

Alerta ID:	119
Fecha del reporte:	01/07/2026
Entidad:	Todas las entidades del ecosistema digital
Título:	Código malicioso generado mediante Inteligencia Artificial abusa de funcionalidades nativas de navegadores web para implementar técnicas de Browser Ransomware y bloquear la interacción del usuario
Herramienta de detección	N/A
Activo involucrado:	Estaciones de trabajo Windows, Linux y macOS, navegadores Chromium (Google Chrome, Microsoft Edge, Brave, Opera y Vivaldi), Mozilla Firefox, perfiles de usuario del navegador, aplicaciones web institucionales, portales de historia clínica electrónica, plataformas de telemedicina, aplicaciones SaaS, sistemas de autenticación corporativa, servicios de correo electrónico institucional, estaciones de trabajo administrativas y clínicas, equipos utilizados para acceso remoto y plataformas web empleadas para procesos asistenciales y administrativos
Tipo de alerta:	Alerta
Nivel de riesgo:	Alto

Objetivo:

Informar a las organizaciones del sector salud en Colombia sobre una técnica emergente de **Browser Ransomware**, divulgada recientemente por investigadores de seguridad, que demuestra cómo actores maliciosos pueden abusar de funcionalidades legítimas implementadas en los navegadores web modernos para bloquear la interacción del usuario, simular un proceso de cifrado de información y desplegar mensajes fraudulentos de extorsión sin comprometer directamente el sistema operativo ni cifrar archivos almacenados en el dispositivo.

La investigación evidencia que herramientas de Inteligencia Artificial generativa son capaces de producir automáticamente código JavaScript funcional que integra múltiples interfaces de

programación (Web APIs) disponibles en navegadores modernos, incluyendo **Fullscreen API**, **Pointer Lock API**, **History API**, manipulación dinámica del **Document Object Model (DOM)** y eventos asociados a dispositivos de entrada, permitiendo construir aplicaciones web que restringen significativamente la capacidad del usuario para abandonar la página o recuperar el control de la sesión de navegación.

A diferencia del ransomware tradicional, cuyo propósito consiste en cifrar archivos mediante algoritmos criptográficos y exigir posteriormente un pago para su recuperación, esta técnica explota mecanismos legítimos del navegador para inducir a la víctima a creer que el equipo ha sido comprometido. La combinación de ingeniería social, manipulación de la interfaz gráfica y abuso de funcionalidades nativas permite generar una percepción altamente convincente de pérdida del control del dispositivo, incrementando la probabilidad de que el usuario siga instrucciones fraudulentas o facilite información sensible.

Aunque la prueba de concepto presentada corresponde a una demostración técnica, las capacidades observadas representan un escenario de riesgo relevante debido a que evidencian cómo la Inteligencia Artificial puede reducir considerablemente la barrera técnica para desarrollar código malicioso orientado a navegadores web, facilitando la creación de nuevas variantes de amenazas que podrían incorporar funcionalidades adicionales como robo de credenciales, distribución de malware, secuestro de sesiones autenticadas o fraude financiero. El propósito de esta alerta es que las entidades evalúen el impacto potencial que este tipo de amenazas podría generar sobre:

- estaciones de trabajo administrativas;
- equipos utilizados para procesos clínicos;
- navegadores corporativos utilizados para acceder a aplicaciones institucionales;
- plataformas de Historia Clínica Electrónica (HCE);
- aplicaciones de telemedicina;
- plataformas de autenticación corporativa;
- sistemas de gestión documental;
- aplicaciones SaaS utilizadas por la organización;
- servicios de correo electrónico institucional;
- activos críticos cuya operación dependa de aplicaciones web.

En el sector salud, donde una parte significativa de los procesos asistenciales y administrativos depende del acceso permanente a plataformas web institucionales, el abuso de funcionalidades legítimas implementadas en los navegadores modernos puede generar interrupciones operativas, afectar la disponibilidad de servicios, inducir al personal a ejecutar acciones inseguras y facilitar la ejecución de campañas posteriores de ingeniería social o compromiso de credenciales institucionales.

Descripción:

Investigadores de seguridad presentaron recientemente una demostración técnica que evidencia la viabilidad de desarrollar código malicioso mediante herramientas de Inteligencia Artificial generativa para implementar una modalidad de **Browser Ransomware**, basada en el abuso de funcionalidades legítimas incorporadas en los navegadores web modernos. La investigación demuestra que asistentes de Inteligencia Artificial son capaces de generar, a partir de instrucciones relativamente simples, aplicaciones web que integran múltiples APIs del navegador con el propósito de restringir progresivamente la interacción del usuario y simular el comportamiento visual de un ataque de ransomware.

A diferencia de campañas convencionales de ransomware, en las cuales el atacante obtiene ejecución de código sobre el sistema operativo para cifrar archivos, modificar configuraciones o establecer persistencia, esta técnica permanece completamente dentro del contexto de ejecución del navegador. El código malicioso se ejecuta como una aplicación web y aprovecha funcionalidades disponibles conforme a los estándares definidos para la plataforma web, sin requerir la explotación de vulnerabilidades conocidas ni la instalación previa de software malicioso sobre el equipo comprometido.

Desde una perspectiva técnica, el ataque combina diversas capacidades expuestas por las Web APIs modernas. Entre ellas destacan la **Fullscreen API**, utilizada para ocultar los elementos de la interfaz del navegador y proporcionar una experiencia inmersiva; la **Pointer Lock API**, que restringe temporalmente el comportamiento del puntero dentro de la página; la manipulación del **History API**, empleada para dificultar la navegación hacia páginas anteriores; y la modificación dinámica del **Document Object Model (DOM)**, que permite actualizar continuamente la interfaz para mostrar mensajes intimidatorios, temporizadores, advertencias de supuesto cifrado y solicitudes fraudulentas de pago.

La investigación pone de manifiesto un cambio significativo en la evolución de las amenazas basadas en navegadores. En lugar de depender exclusivamente de conocimientos avanzados de programación, los actores maliciosos pueden apoyarse en modelos de Inteligencia Artificial capaces de producir código JavaScript funcional, optimizar la interacción entre diferentes APIs del navegador y generar interfaces gráficas altamente convincentes en cuestión de minutos. Como consecuencia, disminuye considerablemente el esfuerzo técnico necesario para desarrollar pruebas de concepto que posteriormente podrían evolucionar hacia campañas maliciosas de mayor complejidad.

Desde el punto de vista defensivo, esta demostración confirma que la superficie de ataque de las aplicaciones web continúa ampliándose a medida que los navegadores incorporan nuevas funcionalidades orientadas a mejorar la experiencia del usuario. Aunque dichas capacidades fueron diseñadas con fines legítimos, su utilización coordinada puede ser aprovechada para implementar mecanismos avanzados de engaño visual, dificultar la recuperación de la sesión de navegación y aumentar la efectividad de campañas de ingeniería social dirigidas contra usuarios con distintos niveles de conocimiento técnico.

En organizaciones del sector salud, donde los profesionales acceden diariamente a historias clínicas electrónicas, plataformas de telemedicina, aplicaciones administrativas, sistemas financieros y servicios de autenticación mediante navegadores web, este tipo de técnicas representa un riesgo creciente. Un usuario que interprete erróneamente que su estación de trabajo ha sido comprometida por un ransomware podría interrumpir procesos asistenciales, reportar incidentes inexistentes, abandonar procedimientos críticos o seguir instrucciones proporcionadas por actores maliciosos durante fases posteriores de la campaña.

Adicionalmente, aunque la prueba de concepto no implementa capacidades reales de cifrado de información ni mecanismos de persistencia sobre el sistema operativo, las técnicas demostradas pueden integrarse fácilmente con campañas de phishing, robo de credenciales, distribución de malware, secuestro de sesiones autenticadas o descarga de componentes maliciosos adicionales, incrementando significativamente el impacto potencial sobre infraestructuras tecnológicas críticas del sector salud.

Consecuencias de la explotación



La explotación exitosa de esta técnica de **Browser Ransomware** puede generar impactos significativos sobre la disponibilidad operativa, la integridad de los procesos institucionales y la confianza de los usuarios en las plataformas tecnológicas utilizadas por la organización. Aunque la prueba de concepto presentada por los investigadores no implementa mecanismos reales de cifrado de archivos ni compromete directamente el sistema operativo, el abuso coordinado de funcionalidades nativas del navegador puede provocar interrupciones operativas, afectar la continuidad de los servicios y facilitar la ejecución de fases posteriores de un ataque más complejo.

Uno de los principales riesgos corresponde a la manipulación de la percepción del usuario mediante el bloqueo parcial o total de la interfaz del navegador. La utilización combinada de APIs como **Fullscreen API**, **Pointer Lock API**, **History API** y la manipulación dinámica del **Document Object Model (DOM)** permite ocultar los elementos propios del navegador, restringir las opciones de navegación y presentar interfaces gráficas que simulan procesos de cifrado, temporizadores de cuenta regresiva y mensajes de extorsión con un elevado nivel de realismo. Como consecuencia, la víctima puede interpretar erróneamente que el sistema operativo ha sido comprometido por un ransomware real.

La capacidad de restringir la interacción del usuario con la sesión de navegación constituye otro de los aspectos más relevantes de esta técnica. Dependiendo del navegador utilizado y de las funcionalidades soportadas, el código malicioso puede dificultar el cierre de la pestaña, interceptar determinados eventos asociados al teclado y al puntero, generar ventanas emergentes repetitivas, modificar continuamente el historial de navegación y reconstruir dinámicamente la interfaz cuando el usuario intenta abandonar la página. Aunque estos mecanismos no impiden completamente recuperar el control del equipo, incrementan significativamente el tiempo requerido para identificar la naturaleza del incidente y favorecen el éxito de la ingeniería social.

Desde una perspectiva operacional, este tipo de ataques puede afectar la disponibilidad de aplicaciones institucionales que dependen completamente del navegador para su funcionamiento. En organizaciones del sector salud, donde múltiples procesos clínicos y administrativos son ejecutados mediante plataformas web, una falsa percepción de

compromiso puede ocasionar la suspensión temporal de actividades asistenciales, el cierre preventivo de estaciones de trabajo, la interrupción de procesos de atención o la activación innecesaria de procedimientos internos de respuesta a incidentes, generando un impacto operativo desproporcionado frente a la naturaleza real del ataque.

Otro riesgo significativo corresponde al aprovechamiento de esta técnica como mecanismo de acceso inicial dentro de campañas maliciosas más amplias. Debido a que el Browser Ransomware centra la atención del usuario en una supuesta infección crítica del equipo, los actores de amenaza disponen de una oportunidad para inducir a la víctima a ejecutar acciones adicionales, tales como descargar herramientas de soporte fraudulentas, instalar aplicaciones aparentemente destinadas a recuperar el sistema, proporcionar credenciales corporativas, autorizar conexiones remotas o realizar pagos asociados a supuestos procesos de recuperación de la información.

La utilización de código generado mediante Inteligencia Artificial introduce además un escenario de riesgo adicional relacionado con la proliferación de amenazas. La automatización del desarrollo de aplicaciones maliciosas permite que actores con conocimientos limitados produzcan rápidamente nuevas variantes capaces de modificar su comportamiento, incorporar funcionalidades adicionales o adaptar la interfaz fraudulenta a diferentes idiomas, sectores económicos o plataformas tecnológicas, incrementando considerablemente la superficie de exposición de las organizaciones.

En entornos hospitalarios, donde las estaciones de trabajo suelen mantener acceso simultáneo a historias clínicas electrónicas, plataformas de telemedicina, sistemas administrativos, aplicaciones financieras y servicios de autenticación corporativa, un incidente de estas características puede generar confusión entre el personal asistencial y administrativo, retrasar la atención de pacientes, afectar procesos críticos y aumentar la carga operativa de los equipos responsables de ciberseguridad y soporte tecnológico.

Adicionalmente, el abuso de funcionalidades legítimas implementadas por los navegadores modernos dificulta la identificación temprana de este tipo de amenazas mediante controles tradicionales basados exclusivamente en firmas o indicadores asociados a malware convencional. Al ejecutarse completamente dentro del contexto del navegador y no requerir explotación de vulnerabilidades conocidas ni instalación de componentes persistentes sobre el sistema operativo, la actividad puede pasar inadvertida para algunos mecanismos de detección

que no realizan inspección del comportamiento de aplicaciones web o análisis contextual de la interacción del usuario.

En escenarios donde esta técnica sea integrada con campañas de phishing o distribución de malware, el Browser Ransomware podría convertirse en una fase inicial destinada a distraer al usuario mientras se ejecutan otras actividades maliciosas, incluyendo el robo de credenciales, el secuestro de sesiones autenticadas, la descarga de cargas útiles adicionales o la instalación de herramientas de acceso remoto. Esta convergencia entre ingeniería social, abuso de APIs del navegador e Inteligencia Artificial representa una evolución significativa en las técnicas empleadas por actores de amenaza para comprometer usuarios y organizaciones.

En escenarios de explotación exitosa, un atacante podría:

- bloquear visualmente la interfaz del navegador mediante el abuso de la **Fullscreen API**;
- restringir parcialmente la interacción del usuario utilizando funcionalidades como **Pointer Lock API** y eventos del teclado;
- manipular dinámicamente el **Document Object Model (DOM)** para simular procesos de cifrado y mensajes de extorsión;
- dificultar la navegación mediante modificaciones sobre el historial de la sesión utilizando **History API**;
- generar una falsa percepción de compromiso del sistema operativo para inducir al pago de rescates inexistentes;
- ejecutar campañas avanzadas de ingeniería social orientadas al robo de credenciales o instalación de software adicional;
- interrumpir temporalmente procesos clínicos y administrativos dependientes de aplicaciones web;
- incrementar la carga operativa de los equipos de respuesta a incidentes debido a falsas alarmas de ransomware;
- facilitar la distribución de componentes maliciosos adicionales durante el proceso de engaño al usuario;
- utilizar aplicaciones web aparentemente legítimas como plataforma para desarrollar campañas posteriores dirigidas contra la infraestructura tecnológica de la organización.

Aunque la demostración técnica presentada no incorpora capacidades de cifrado de información ni persistencia sobre el sistema operativo, las técnicas observadas evidencian

cómo el abuso coordinado de funcionalidades legítimas del navegador, combinado con herramientas de Inteligencia Artificial generativa, puede evolucionar rápidamente hacia campañas más sofisticadas que integren robo de información, secuestro de sesiones autenticadas, distribución de malware, fraude financiero o compromiso de aplicaciones críticas utilizadas por organizaciones del sector salud.

Modo de explotación del ataque



Fase 1 – Desarrollo del código malicioso mediante Inteligencia Artificial generativa

La cadena de ataque inicia con la utilización de plataformas de Inteligencia Artificial generativa capaces de producir automáticamente aplicaciones web desarrolladas en HTML5, CSS y JavaScript a partir de instrucciones proporcionadas por el operador. A diferencia del desarrollo tradicional de software malicioso, donde el atacante requiere conocimientos especializados sobre programación web, modelos de interacción del navegador y manipulación de eventos, la Inteligencia Artificial reduce significativamente la complejidad técnica necesaria para generar una prueba de concepto funcional.

Durante esta fase, el actor de amenaza define los objetivos operativos del código malicioso mediante instrucciones orientadas a implementar un falso ransomware basado en navegador. Como resultado, el modelo genera automáticamente componentes capaces de manipular la interfaz gráfica, controlar el comportamiento de la sesión de navegación, interceptar eventos asociados a dispositivos de entrada y presentar mensajes visuales diseñados para simular un incidente crítico de seguridad.

Aunque la demostración presentada corresponde a un entorno controlado de investigación, el mismo enfoque podría ser utilizado para desarrollar variantes orientadas al robo de credenciales, distribución de malware, secuestro de sesiones autenticadas o campañas avanzadas de ingeniería social.

Las actividades realizadas durante esta fase pueden incluir:

- definición de los objetivos funcionales del código malicioso;

- generación automática de código HTML, CSS y JavaScript mediante herramientas de Inteligencia Artificial;
- integración de APIs nativas del navegador;
- construcción de interfaces gráficas orientadas a simular ransomware;
- desarrollo automatizado de funciones para manipular la interacción del usuario;
- optimización del código generado para diferentes navegadores;
- preparación de la aplicación web para su distribución.

El objetivo principal consiste en reducir el tiempo de desarrollo del código malicioso y disminuir la barrera técnica necesaria para construir aplicaciones web con capacidades avanzadas de manipulación del navegador.

Fase 2 – Distribución de la aplicación web maliciosa e interacción inicial con la víctima

Una vez desarrollado el código, los actores de amenaza proceden a distribuir la aplicación web mediante técnicas tradicionales de ingeniería social, campañas de phishing, anuncios patrocinados (malvertising), publicaciones en redes sociales, mensajería instantánea o sitios web comprometidos. En escenarios más sofisticados, la aplicación puede alojarse en dominios que simulan corresponder a organizaciones legítimas o utilizar certificados TLS válidos con el propósito de incrementar la percepción de confianza durante el acceso inicial.

El ataque depende de que la víctima acceda voluntariamente al sitio web malicioso utilizando un navegador compatible con las funcionalidades requeridas por la aplicación. Debido a que la ejecución ocurre completamente dentro del contexto del navegador, no resulta necesaria la instalación previa de software ni la explotación de vulnerabilidades sobre el sistema operativo. Durante esta etapa, la aplicación identifica automáticamente las capacidades soportadas por el navegador, valida la disponibilidad de las APIs necesarias y prepara el entorno para iniciar la simulación del ransomware.

Las actividades realizadas durante esta fase pueden incluir:

- distribución del enlace malicioso mediante campañas de ingeniería social;
- publicación de la aplicación en sitios aparentemente legítimos;
- utilización de dominios con apariencia confiable;

- acceso inicial del usuario al sitio web;
- identificación del navegador utilizado;
- validación de las APIs disponibles;
- preparación del entorno para la ejecución del ataque.

El objetivo consiste en lograr que la víctima ejecute voluntariamente la aplicación web dentro de un navegador compatible con las funcionalidades requeridas por el código malicioso.

Fase 3 – Abuso de APIs nativas del navegador y bloqueo de la interacción del usuario

Después de verificar que el entorno reúne las condiciones necesarias, la aplicación comienza a utilizar de forma coordinada diversas interfaces de programación implementadas por el navegador para restringir progresivamente la capacidad del usuario de controlar la sesión de navegación.

Entre las principales funcionalidades utilizadas se encuentran la Fullscreen API, empleada para ocultar la interfaz propia del navegador; la Pointer Lock API, utilizada para limitar temporalmente el comportamiento del puntero; la History API, que dificulta la navegación hacia páginas anteriores; y diversos eventos asociados al teclado y al mouse que permiten controlar parcialmente la interacción del usuario con la aplicación.

El empleo simultáneo de estas capacidades produce una experiencia visual inmersiva que dificulta distinguir entre la interfaz del navegador y la simulación presentada por el código malicioso, incrementando considerablemente la credibilidad del supuesto incidente de seguridad.

Las actividades realizadas durante esta fase pueden incluir:

- activación del modo de pantalla completa;
- ocultamiento de elementos de navegación del navegador;
- restricción parcial del movimiento del puntero;
- manipulación del historial de navegación;
- interceptación de eventos asociados al teclado;
- control de eventos del mouse;

- preparación del entorno visual para la simulación del ransomware.

El objetivo consiste en reducir la capacidad del usuario para abandonar la sesión y generar la percepción de que el equipo ha perdido completamente el control operativo.

Fase 4 – Simulación del proceso de ransomware mediante manipulación dinámica del DOM

Una vez restringida la interacción del usuario, la aplicación inicia la representación visual del supuesto ataque utilizando técnicas avanzadas de manipulación dinámica del Document Object Model (DOM).

La interfaz puede mostrar mensajes que indican un presunto cifrado de archivos, temporizadores de cuenta regresiva, advertencias sobre pérdida permanente de información, solicitudes de pago en criptomonedas y otros elementos característicos de campañas reales de ransomware. Todos estos componentes son generados dinámicamente mediante JavaScript sin que exista realmente un proceso de cifrado ejecutándose sobre el sistema operativo.

La actualización permanente del contenido del DOM permite modificar la interfaz en tiempo real, presentar animaciones, simular actividades de cifrado y responder a determinadas acciones realizadas por el usuario durante la interacción con la página.

Las actividades realizadas durante esta fase pueden incluir:

- generación dinámica de mensajes de extorsión;
- simulación visual del cifrado de archivos;
- despliegue de temporizadores falsos;
- actualización continua de la interfaz gráfica;
- modificación dinámica del contenido HTML;
- incorporación de elementos audiovisuales destinados a incrementar el impacto psicológico;
- presentación de instrucciones fraudulentas para el supuesto proceso de recuperación.

El objetivo consiste en convencer a la víctima de que el dispositivo ha sido comprometido por un ransomware legítimo y aumentar la probabilidad de que siga las instrucciones proporcionadas por los atacantes.

Fase 5 – Ingeniería social y manipulación psicológica del usuario

Con la simulación completamente desplegada, el ataque se centra en explotar la reacción emocional de la víctima. El código malicioso presenta mensajes diseñados para generar sensación de urgencia, pérdida inminente de información y presión psicológica, induciendo al usuario a actuar sin verificar la autenticidad del incidente.

Dependiendo de la campaña implementada, la interfaz puede solicitar pagos, credenciales corporativas, información personal, instalación de aplicaciones de recuperación o contacto con supuestos servicios de soporte técnico controlados por los actores de amenaza.

La efectividad de esta fase depende principalmente de la credibilidad alcanzada durante las etapas anteriores y del desconocimiento del usuario respecto al funcionamiento interno del navegador.

Las actividades realizadas durante esta fase pueden incluir:

- despliegue de mensajes intimidatorios;
- simulación de pérdida irreversible de información;
- generación de temporizadores para incrementar la presión psicológica;
- solicitud fraudulenta de pagos;
- inducción al ingreso de credenciales;
- redirección hacia portales controlados por el atacante;
- manipulación del comportamiento del usuario mediante ingeniería social.

El objetivo consiste en obtener un beneficio económico o facilitar fases posteriores de compromiso aprovechando la percepción de un incidente crítico inexistente.

Fase 6 – Evasión de controles tradicionales y permanencia de la sesión fraudulenta

Durante toda la ejecución del ataque, la aplicación procura mantener la apariencia de legitimidad utilizando exclusivamente funcionalidades soportadas por el navegador. Debido a que el código opera dentro del contexto de una aplicación web y no instala componentes persistentes sobre el sistema operativo, muchos controles tradicionales orientados a detectar malware convencional pueden no generar alertas inmediatas.

El comportamiento del código se limita al uso intensivo de APIs autorizadas por el navegador, reduciendo la necesidad de explotar vulnerabilidades o ejecutar procesos externos. Esta característica incrementa la dificultad para diferenciar entre una aplicación web legítima y una diseñada con fines maliciosos únicamente mediante mecanismos de inspección basados en firmas.

Las actividades realizadas durante esta fase pueden incluir:

- utilización exclusiva de APIs nativas del navegador;
- ejecución íntegra dentro del contexto web;
- ausencia de modificaciones sobre el sistema operativo;
- mantenimiento del modo de pantalla completa;
- reconstrucción dinámica de la interfaz durante la sesión;
- reducción de indicadores tradicionales de malware;
- prolongación de la interacción fraudulenta con la víctima.

El objetivo consiste en mantener activa la simulación durante el mayor tiempo posible y dificultar que el usuario identifique la verdadera naturaleza del incidente.

Fase 7 – Monetización y posible evolución hacia campañas de compromiso adicionales

Una vez obtenida la respuesta esperada por parte de la víctima, los actores de amenaza pueden proceder a monetizar el ataque mediante diferentes mecanismos, incluyendo solicitudes fraudulentas de pago, robo de credenciales, distribución de malware adicional o instalación de herramientas de acceso remoto.

Aunque la demostración técnica presentada por los investigadores se limita a evidenciar la viabilidad del Browser Ransomware, la misma arquitectura podría evolucionar fácilmente hacia campañas híbridas que combinen ingeniería social, robo de identidad, secuestro de sesiones autenticadas, fraude financiero y distribución de código malicioso adicional.

En organizaciones del sector salud, este escenario incrementa significativamente el riesgo de compromiso de credenciales institucionales, acceso no autorizado a aplicaciones clínicas, interrupción de procesos asistenciales y utilización de estaciones de trabajo comprometidas como punto inicial para ataques dirigidos contra otros activos tecnológicos de la organización. Las actividades realizadas durante esta fase pueden incluir:

- obtención fraudulenta de pagos;
- captura de credenciales institucionales;
- distribución de componentes maliciosos adicionales;
- instalación de herramientas de acceso remoto;
- secuestro de sesiones autenticadas;
- utilización de la estación comprometida para campañas posteriores;
- ampliación del compromiso hacia otros activos tecnológicos de la organización.

El objetivo consiste en maximizar el beneficio obtenido mediante la ingeniería social y aprovechar la confianza depositada por la víctima para facilitar ataques posteriores con un mayor impacto sobre la infraestructura tecnológica.

Vulnerabilidades asociadas



Abuso de funcionalidades legítimas del navegador (Web APIs)

- **Descripción:** La técnica demostrada no explota vulnerabilidades tradicionales del sistema operativo ni fallos de seguridad del navegador. En su lugar, utiliza de forma coordinada diversas Web APIs implementadas por los navegadores modernos, entre ellas Fullscreen API, Pointer Lock API, History API, Keyboard Events y la manipulación dinámica del Document Object Model (DOM). Estas funcionalidades fueron diseñadas para proporcionar capacidades avanzadas a aplicaciones web legítimas; sin embargo, cuando son utilizadas de manera maliciosa pueden restringir la interacción del usuario,

ocultar elementos de la interfaz del navegador y crear escenarios altamente convincentes de ingeniería social.

- **Impacto:** El abuso de funcionalidades legítimas permite construir aplicaciones web que dificultan la recuperación de la sesión de navegación, generan falsas percepciones de compromiso del equipo y facilitan campañas de fraude, phishing, robo de credenciales y distribución de malware sin necesidad de explotar vulnerabilidades conocidas.
 - **Referencias técnicas:**
 - MITRE ATT&CK: T1056 – Input Capture; T1566 – Phishing; T1204 – User Execution.
 - CWE: CWE-451 – User Interface (UI) Misrepresentation of Critical Information; CWE-1021 – Improper Restriction of Rendered UI Layers or Frames.
-

Ingeniería social mediante simulación de ransomware

- **Descripción:** El Browser Ransomware basa su efectividad en la manipulación psicológica del usuario. Mediante interfaces gráficas diseñadas para simular procesos de cifrado, mensajes de extorsión y temporizadores de cuenta regresiva, los atacantes buscan inducir a la víctima a creer que el dispositivo ha sido comprometido por un ransomware legítimo. Este enfoque explota la confianza del usuario en la apariencia visual de la interfaz y en el desconocimiento del funcionamiento interno del navegador.
 - **Impacto:** La explotación exitosa puede conducir a pagos fraudulentos, divulgación de información sensible, instalación voluntaria de software malicioso, entrega de credenciales corporativas y activación innecesaria de procedimientos de respuesta a incidentes.
 - **Referencias técnicas:**
 - MITRE ATT&CK: T1656 – Impersonation; T1566 – Phishing; T1204.001 – User Execution: Malicious Link.
 - CWE: CWE-451 – User Interface (UI) Misrepresentation of Critical Information.
-

Ejecución de código JavaScript no confiable

- **Descripción:** La totalidad del ataque se desarrolla mediante código JavaScript ejecutado dentro del contexto del navegador. Cuando el usuario accede a un sitio web controlado por un actor malicioso, el navegador interpreta automáticamente el código suministrado por la aplicación web, permitiendo que éste invoque diversas APIs autorizadas por la plataforma. Aunque el modelo de seguridad del navegador limita el acceso directo al sistema operativo, la ejecución de código no confiable continúa representando un riesgo significativo cuando se combina con técnicas de ingeniería social y abuso de funcionalidades legítimas.
- **Impacto:** La ejecución de código JavaScript malicioso puede facilitar el robo de credenciales, manipulación de formularios, secuestro de sesiones, redireccionamientos fraudulentos, distribución de malware y ejecución de ataques posteriores orientados al compromiso de la infraestructura tecnológica.
- **Referencias técnicas:**
 - MITRE ATT&CK: T1059.007 – JavaScript; T1189 – Drive-by Compromise.
 - CWE: CWE-79 – Improper Neutralization of Input During Web Page Generation (Cross-site Scripting); CWE-95 – Improper Neutralization of Directives in Dynamically Evaluated Code.

Reducción de la barrera técnica mediante Inteligencia Artificial generativa

- **Descripción:** La investigación demuestra que modelos de Inteligencia Artificial generativa pueden producir automáticamente aplicaciones web funcionales a partir de instrucciones en lenguaje natural. Esta capacidad disminuye considerablemente el conocimiento técnico requerido para desarrollar herramientas ofensivas, acelerando el ciclo de creación de nuevas variantes de código malicioso y facilitando que actores con experiencia limitada desarrollen pruebas de concepto cada vez más sofisticadas.
- **Impacto:** La automatización del desarrollo incrementa el volumen potencial de amenazas, acelera la evolución de campañas maliciosas y dificulta la anticipación de nuevas variantes por parte de los equipos de defensa.

- **Referencias técnicas:**

- MITRE ATT&CK: No aplica directamente (facilitador del desarrollo de capacidades ofensivas).
- CWE: No aplica directamente.

Dependencia de la interacción del usuario

- **Descripción:** La técnica requiere que la víctima acceda voluntariamente al sitio web y otorgue determinadas interacciones iniciales, como hacer clic sobre la página para habilitar algunas APIs del navegador o aceptar solicitudes relacionadas con el modo de pantalla completa. Esta dependencia convierte al usuario en un componente crítico de la cadena de ataque y evidencia la importancia de la concienciación en ciberseguridad como mecanismo de defensa.
- **Impacto:** La interacción del usuario puede habilitar la ejecución de funcionalidades avanzadas del navegador que posteriormente serán utilizadas para desarrollar campañas de fraude, phishing, secuestro de sesiones o distribución de software malicioso.
- **Referencias técnicas:**
 - MITRE ATT&CK: T1204 – User Execution; T1566 – Phishing.
 - CWE: CWE-640 – Weak Password Recovery Mechanism for Forgotten Password (ingeniería social relacionada); CWE-451 – User Interface (UI) Misrepresentation of Critical Information.

Ausencia de controles de endurecimiento del navegador

- **Descripción:** La explotación de esta técnica resulta más efectiva en entornos donde los navegadores mantienen habilitadas por defecto todas las funcionalidades relacionadas con APIs de interacción avanzada y donde no existen políticas corporativas para restringir el uso de pantalla completa, controlar extensiones, limitar permisos de sitios web o supervisar el comportamiento de aplicaciones web de alto riesgo. La ausencia de

mecanismos de hardening incrementa la superficie de ataque y facilita el abuso de funcionalidades legítimas.

- **Impacto:** Configuraciones inseguras del navegador pueden favorecer la ejecución de aplicaciones web maliciosas, incrementar la exposición frente a campañas de ingeniería social y dificultar la detección temprana de comportamientos anómalos.
 - **Referencias técnicas:**
 - MITRE ATT&CK: T1580 – Cloud Infrastructure Discovery (contexto de reconocimiento de capacidades); T1204 – User Execution.
 - CWE: CWE-16 – Configuration.
-

Deficiencias en la concienciación sobre amenazas basadas en navegadores

- **Descripción:** Las campañas de Browser Ransomware explotan principalmente el desconocimiento de los usuarios sobre el funcionamiento de los navegadores modernos y las diferencias existentes entre una aplicación web maliciosa y un ransomware real. La ausencia de programas continuos de sensibilización incrementa la probabilidad de que los usuarios interpreten erróneamente la situación y ejecuten acciones solicitadas por los atacantes.
- **Impacto:** La falta de concienciación favorece el éxito de campañas de ingeniería social, incrementa el riesgo de compromiso de credenciales institucionales y puede ocasionar interrupciones operativas innecesarias dentro de la organización.
- **Referencias técnicas:**
 - MITRE ATT&CK: T1566 – Phishing; T1656 – Impersonation.
 - CWE: CWE-451 – User Interface (UI) Misrepresentation of Critical Information; CWE-200 – Exposure of Sensitive Information to an Unauthorized Actor (como consecuencia de la ingeniería social).

Recomendaciones

Las organizaciones del sector salud deben fortalecer los controles de seguridad implementados sobre las estaciones de trabajo y los navegadores corporativos con el propósito de reducir el riesgo asociado al abuso de funcionalidades legítimas de las aplicaciones web. Aunque la técnica demostrada no explota vulnerabilidades del sistema operativo, evidencia cómo el uso coordinado de APIs nativas del navegador puede facilitar campañas avanzadas de ingeniería social, afectar la continuidad operativa y servir como mecanismo de acceso inicial para ataques de mayor complejidad.

Se recomienda implementar las siguientes medidas preventivas:

Fortalecer la configuración de seguridad de los navegadores corporativos

Implementar políticas corporativas de endurecimiento (Hardening) para Google Chrome, Microsoft Edge, Mozilla Firefox y demás navegadores autorizados por la organización. Estas políticas deben restringir, cuando sea técnicamente viable, el uso innecesario de funcionalidades como Fullscreen API, controlar la instalación de extensiones, limitar permisos concedidos a sitios web y aplicar configuraciones centralizadas mediante políticas de grupo (GPO), Microsoft Intune, Google Workspace u otras plataformas de administración empresarial.

Restringir el acceso a sitios web no autorizados

Implementar mecanismos de filtrado web que permitan bloquear el acceso a dominios recientemente registrados, sitios con baja reputación, servicios de alojamiento utilizados para distribuir campañas maliciosas y páginas que no sean necesarias para el desarrollo de las funciones institucionales. Complementariamente, mantener actualizadas las categorías de filtrado implementadas por Secure Web Gateway (SWG), Firewalls de Nueva Generación (NGFW) o plataformas Secure Service Edge (SSE).

Mantener permanentemente actualizados los navegadores

Garantizar que todos los navegadores utilizados dentro de la organización se encuentren actualizados a las versiones más recientes suministradas por sus fabricantes. Las actualizaciones incorporan mejoras relacionadas con controles de seguridad, restricciones sobre determinadas APIs, mecanismos de aislamiento de procesos y corrección de vulnerabilidades que podrían ser utilizadas en campañas más sofisticadas.

Supervisar continuamente la actividad de los navegadores

Incorporar eventos relacionados con navegadores dentro de las plataformas SIEM para facilitar la detección temprana de comportamientos inusuales. Se recomienda correlacionar registros provenientes del navegador, del sistema operativo, del proxy web, del firewall, del DNS corporativo y de soluciones EDR con el fin de identificar accesos a dominios sospechosos, múltiples solicitudes de pantalla completa, navegación hacia sitios de baja reputación o patrones asociados con campañas de Browser Ransomware.

Aplicar el principio de mínimo privilegio

Restringir el uso de cuentas con privilegios administrativos sobre estaciones de trabajo utilizadas para navegación web. Los usuarios deben operar con privilegios estándar, limitando la capacidad de instalar software, modificar configuraciones críticas del sistema o alterar políticas corporativas del navegador.

Fortalecer la concienciación del personal

Desarrollar campañas periódicas de sensibilización orientadas a explicar el funcionamiento de amenazas basadas en navegadores, enfatizando que la aparición de una pantalla de supuesto ransomware no implica necesariamente el compromiso del sistema operativo. El personal debe conocer los procedimientos institucionales para reportar este tipo de incidentes y evitar realizar pagos, proporcionar credenciales o instalar software recomendado por sitios web desconocidos.

Implementar controles de aislamiento del navegador

Cuando la arquitectura tecnológica lo permita, evaluar la implementación de soluciones de Remote Browser Isolation (RBI) o mecanismos equivalentes que ejecuten las sesiones de navegación en entornos aislados, reduciendo el impacto de aplicaciones web potencialmente maliciosas sobre las estaciones de trabajo institucionales.

Fortalecer los controles de acceso a aplicaciones web críticas

Implementar autenticación multifactor (MFA) en todas las aplicaciones institucionales accesibles mediante navegador, especialmente aquellas relacionadas con Historia Clínica Electrónica, plataformas financieras, telemedicina, correo electrónico, directorios corporativos y servicios en la nube. De esta manera se reduce el impacto potencial en caso de que campañas de Browser Ransomware evolucionen hacia el robo de credenciales.

Implementar controles de seguridad sobre JavaScript y contenido activo

Evaluar la utilización de mecanismos que permitan controlar la ejecución de contenido activo en aplicaciones web, tales como Content Security Policy (CSP), listas de permitidos (Allow Lists), controles de reputación web y políticas que restrinjan la ejecución de scripts provenientes de orígenes no confiables, reduciendo la posibilidad de abuso de funcionalidades legítimas del navegador.

Establecer procedimientos de respuesta para incidentes relacionados con navegadores

Actualizar los procedimientos institucionales de respuesta a incidentes incorporando escenarios específicos relacionados con Browser Ransomware, phishing avanzado y abuso de Web APIs. Los equipos de soporte y ciberseguridad deben disponer de procedimientos documentados para recuperar rápidamente el control de la sesión del navegador, verificar la

ausencia de compromiso del sistema operativo y validar la integridad de las aplicaciones institucionales antes de restablecer la operación.

Realizar ejercicios de simulación y pruebas de concienciación

Desarrollar ejercicios periódicos que permitan evaluar la capacidad del personal para identificar campañas de ingeniería social basadas en navegadores. Estas actividades deben incluir simulaciones controladas, análisis posterior de resultados y retroalimentación dirigida a fortalecer la cultura de ciberseguridad dentro de la organización.

Fortalecer la arquitectura de Zero Trust para el acceso a aplicaciones web

Adoptar principios de Zero Trust, verificando continuamente la identidad del usuario, el estado de seguridad del dispositivo y el contexto de acceso antes de permitir la interacción con aplicaciones institucionales críticas. La aplicación de acceso condicional, evaluación continua del riesgo y segmentación lógica reduce significativamente el impacto potencial de campañas que intenten utilizar el navegador como vector inicial de compromiso.

La evolución de las amenazas impulsadas por Inteligencia Artificial evidencia que los navegadores web se han convertido en un componente estratégico dentro de la superficie de ataque de las organizaciones. En consecuencia, las entidades del sector salud deben considerar los navegadores corporativos como activos críticos, incorporando controles específicos de endurecimiento, monitoreo continuo, gestión de configuraciones y concienciación del usuario, con el fin de reducir la probabilidad de éxito de campañas que abusen de funcionalidades legítimas de la plataforma web para comprometer la operación institucional.

Conclusiones

La investigación recientemente divulgada evidencia una evolución significativa en la forma en que los actores de amenaza pueden aprovechar las capacidades de la Inteligencia Artificial para desarrollar herramientas ofensivas orientadas a plataformas web modernas. Aunque la prueba de concepto presentada corresponde a un escenario demostrativo y no implementa

procesos reales de cifrado sobre el sistema operativo, demuestra que es posible utilizar exclusivamente funcionalidades legítimas del navegador para generar una percepción altamente convincente de un ataque de ransomware, incrementando la efectividad de campañas de ingeniería social y reduciendo la barrera técnica necesaria para desarrollar este tipo de amenazas.

Desde una perspectiva de ciberseguridad, esta técnica representa un cambio relevante en el paradigma tradicional de las amenazas. En lugar de explotar vulnerabilidades del sistema operativo o instalar código malicioso persistente sobre los equipos comprometidos, el ataque se desarrolla completamente dentro del contexto de ejecución del navegador mediante el uso coordinado de **Fullscreen API**, **Pointer Lock API**, **History API**, manipulación del **Document Object Model (DOM)** y otras funcionalidades implementadas conforme a los estándares de la plataforma web. Este enfoque demuestra que el abuso de características legítimas puede convertirse en un mecanismo eficaz para inducir comportamientos inseguros por parte de los usuarios sin requerir la explotación de vulnerabilidades conocidas.

La incorporación de modelos de Inteligencia Artificial generativa como facilitadores del desarrollo de código ofensivo incrementa considerablemente el riesgo de aparición de nuevas variantes de Browser Ransomware y otras amenazas basadas en navegadores. La capacidad de producir aplicaciones web funcionales mediante instrucciones en lenguaje natural permite acelerar el desarrollo de campañas maliciosas, reducir los tiempos de preparación y facilitar que actores con conocimientos técnicos limitados generen herramientas cada vez más sofisticadas, dificultando la anticipación y respuesta por parte de las organizaciones.

Para el sector salud, este escenario adquiere especial relevancia debido a la creciente dependencia de aplicaciones web para soportar procesos clínicos, administrativos y operacionales. Sistemas de Historia Clínica Electrónica, plataformas de telemedicina, aplicaciones de gestión hospitalaria, portales institucionales, soluciones en la nube y servicios de autenticación son utilizados diariamente mediante navegadores web, convirtiéndolos en componentes críticos cuya indisponibilidad o manipulación puede afectar directamente la continuidad de la prestación de los servicios de salud.

Si bien la técnica analizada no compromete directamente la confidencialidad o integridad de la información almacenada en los equipos, sí puede generar interrupciones operativas, provocar falsas activaciones de procedimientos de respuesta a incidentes, inducir al personal a ejecutar

acciones inseguras y servir como fase inicial para campañas más complejas orientadas al robo de credenciales, distribución de malware, secuestro de sesiones autenticadas o compromiso de infraestructura tecnológica.

En este contexto, resulta fundamental que las organizaciones adopten un enfoque de defensa en profundidad que considere al navegador como un activo crítico dentro de la arquitectura de seguridad. La implementación de controles de endurecimiento, políticas centralizadas de administración de navegadores, monitoreo continuo del comportamiento de aplicaciones web, mecanismos de aislamiento de sesiones, autenticación multifactor, filtrado de contenido y programas permanentes de concienciación permitirá reducir significativamente la superficie de exposición frente a este tipo de amenazas emergentes.

Finalmente, el **CSIRT Salud** recomienda a todas las entidades del sector mantener una vigilancia permanente sobre la evolución de las amenazas asociadas al abuso de funcionalidades legítimas de navegadores web y al uso malicioso de la Inteligencia Artificial. La incorporación temprana de medidas preventivas, capacidades de detección y procedimientos específicos de respuesta contribuirá a fortalecer la resiliencia cibernética institucional y a reducir el impacto potencial que este tipo de técnicas podría generar sobre la disponibilidad, integridad y continuidad de los servicios esenciales del sector salud.

Fuentes:

- <https://thehackernews.com/2026/06/ai-generated-browser-ransomware-abuses.html>
- https://developer.mozilla.org/en-US/docs/Web/API/Pointer_Lock_API
- https://developer.mozilla.org/en-US/docs/Web/API/History_API
- https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model
- <https://developer.mozilla.org/en-US/docs/Web/API/KeyboardEvent>
- <https://attack.mitre.org/>
- <https://cwe.mitre.org/>
- <https://owasp.org/www-project-web-security-testing-guide/>
- <https://owasp.org/www-project-top-ten/>
- <https://developer.mozilla.org/en-US/docs/Web/Security>
- https://developer.mozilla.org/en-US/docs/Web/API/Fullscreen_API/Guide

- <https://chromeenterprise.google/policies/>
- <https://learn.microsoft.com/deployedge/microsoft-edge-security>

En caso de dudas, inquietudes o requerir asistencia adicional relacionada con esta alerta de seguridad, puede comunicarse directamente con el **CSIRT Salud** a través de las líneas telefónicas **(+57) 316 893 1490 - 318 155 3570** o mediante el correo electrónico **csirtsalud@minsalud.gov.co**. Nuestro equipo está disponible para brindar el acompañamiento necesario.