MANUAL CAZADORES

DESINFORMACIÓN Y OPERACIONES DE INFLUENCIA

Cazadores de **Fake News**
cazadoresdefakenews.info

Sobre Cazadores de Fake News

SEPTIEMBRE 2023

En 2019, un grupo de amigos venezolanos nos unimos preocupados por el limitado acceso a Internet y la falta de información fiable en Venezuela. Fundamos una pequeña cibercomunidad con la que intentamos mantenernos informados durante aquella época de largos apagones eléctricos e inestabilidad sociopolítica.

Comenzamos a agruparnos en WhatsApp, compartiendo **artículos e investigaciones** de los principales medios independientes del país a quien quisiera unirse a nuestros grupos. Nos esforzábamos por enviar noticias relevantes y profesionales, firmadas por medios y periodistas en quienes confiamos.

Al poco tiempo, nuestra propia comunidad comenzó a preguntarse cuál era el origen de rumores y posibles «noticias falsas» que recibía a diario en WhatsApp y redes sociales. Y de forma colaborativa, primero en varios grupos de WhatsApp y luego en Telegram, identificamos que existe un problema al que estamos expuestos todos, llamado «desinformación». Y nos animamos a estudiarlo.

Decidimos convertirnos en Cazadores de *fake news* para estudiar cómo se utiliza la desinformación en diferentes contextos políticos y sociales. Desmontamos bulos y rumores de todo tipo y, con el tiempo, comenzamos a investigar operaciones de influencia. Nuestro objetivo es capacitar a la sociedad civil para hacerla resiliente a estos problemas, en la era de la manipulación digital.

¡Únete a nosotros en la lucha contra la desinformación!

Índice

HAZ CLICK EN UN TEMA PARA IR A LA PÁGINA CORRESPONDIENTE

PRIMERA PARTE

Conceptos básicos

01

conceptos básicos

01

PRIMERA PARTE

Conceptos básicos

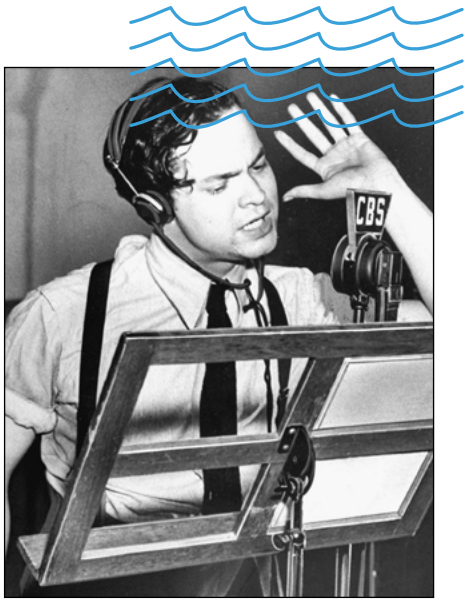
La **alfabetización digital** va más allá de saber cómo usar un dispositivo. Se trata de un proceso de obtención de capacidades que permite que un ciudadano viva, aprenda, trabaje, participe y prospere en una sociedad digital, incluyendo su **capacidad para informarse**.

FUENTE: «A HEALTH AND CARE DIGITAL CAPABILITY FRAMEWORK», NATIONAL HEALTH SERVICE, INGLATERRA (2017), ADAPTACIÓN DE JISC DIGITAL CAPACITY FRAMEWORK (2015)

LA SOCIEDAD CIVIL NECESITA COMPRENDER LAS HERRAMIENTAS TECNOLÓGICAS QUE SE UTILIZAN PARA CONTROLAR Y MANIPULAR LA INFORMACIÓN.

Como algunos países han sido expuestos por años a desinformación en línea, ha sido necesario impulsar en ellos programas de alfabetización digital de la sociedad civil a gran escala. Éstos fomentan el

conceptos básicos



ORSON WELLES ANTE EL MICRÓFONO DURANTE LA EMISIÓN DE LA GUERRA DE LOS MUNDOS EN 1938.
© BRIDGERMAN IMAGES

> LOS SOCIÓLOGOS APUNTAN HOY EN DÍA A QUE EL PODER DE LOS MEDIOS DE INFORMACIÓN CONTRIBUYÓ A CREAR EL MITO POSTERIOR DE QUE UNA GRAN PARTE DE LA POBLACIÓN SE TOMÓ EN SERIO LA INVASIÓN ALIENÍGENA.
>
> NATIONAL GEOGRAPHIC

pensamiento crítico y construyen sociedades más resilientes a operaciones de influencia en línea que, a veces, obedecen a intereses geopolíticos.En un mundo donde la información fluye rápidamente a través de Internet y las redes sociales, es crucial que todos los ciudadanos desarrollen la capacidad de analizar críticamente la información que reciben, para **hacerse menos vulnerables a la manipulación**.

En Cazadores de *fake news* queremos compartir algunos conceptos útiles para **comprender las nuevas dinámicas informativas** y tomar decisiones más acertadas al consumir información.

La defensa de la democracia depende en gran medida de la capacidad que tiene la sociedad civil para navegar en el mundo digital de manera crítica y responsable. La alfabetización digital es la llave que nos permite abrir la puerta hacia un **futuro más informado y consciente**.

Navegando en el mar de la (des)información

Antes de la llegada de internet, la sociedad civil era una receptora de información mucho más pasiva que en la actualidad. Una cantidad menor de medios nos informaron durante décadas a través de la radio, la prensa y la televisión. Publicaron contenidos de altísima calidad y con gran rigurosidad, aunque algunos carecían de esto y no estaban exentos de sesgos editoriales, errores y desinformación.

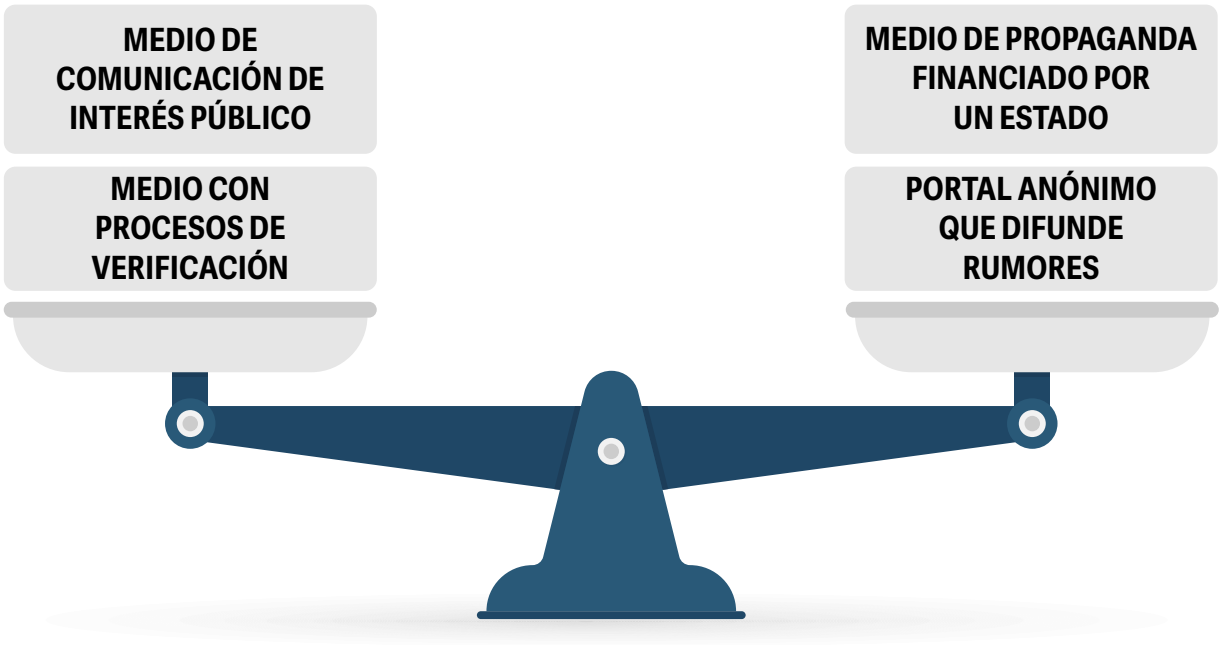
Un ejemplo clásico de «*fake news*» es la famosa transmisión radiofónica de «La guerra de los mundos» realizada por Orson Welles en 1938. En su programa de radio dramatizó un segmento de la historia de H. G. Wells en la que la Tierra era víctima de una invasión alienígena, generando pánico en parte de la audiencia que creyó que la historia era real y convirtiéndose en un **hito del impacto que pueden tener los medios** de comunicación

conceptos básicos

cuando engañan a quien recibe sus contenidos.

En la actualidad, la sociedad civil se encuentra frente a una vasta red de fuentes informativas. Estas incluyen medios de comunicación tradicionales como radio, prensa y televisión, así como redes sociales que, aunque a menudo se consideran medios, son en realidad **canales que transmiten información** generada por diversos actores. Entre estos actores también se encuentran medios, partidos políticos, periodistas, comentaristas, políticos, expertos e influenciadores, quienes, en ocasiones, pueden contribuir a la desinformación.

La relación del ciudadano con la información también se ha complicado debido a una **falsa equivalencia aparente** entre fuentes confiables y no confiables. Aunque ciertos medios parezcan similares en calidad o contenidos, sus intenciones pueden diferir: unos buscan informar con rigor, mientras que otros solo aspiran a persuadir a su público, por distintas razones.



NO PUEDE EQUIPARARSE EL CONTENIDO GENERADO POR MEDIOS QUE BUSCAN INFORMAR CON RIGUROSIDAD Y FUENTES QUE BUSCAN PERSUADIR A FAVOR DE ALGÚN INTERÉS.

conceptos básicos

ES FUNDAMENTAL QUE LA SOCIEDAD CIVIL CONOZCA Y COMPRENDA ESTE PROCESO PARA PODER DESARROLLAR PENSAMIENTO CRÍTICO, LA ÚNICA VACUNA EFECTIVA CONTRA LA DESINFORMACIÓN.

En redes sociales y servicios de mensajería como Telegram existen falsos noticieros que republican información de portales de noticias reales junto con memes, contenido de entretenimiento, publicitario e incluso rumores e información no verificada, con el objetivo de **aumentar sus audiencias o mantener la interacción del público** apelando a sus emociones.

Estar siempre «conectados» y dispuestos a recibir toda la información que los algoritmos de las diferentes redes sociales nos envíen, puede exponernos a gran cantidad de temas sin permitirnos **evaluar detenidamente su veracidad**, priorizando la cantidad sobre la importancia o relevancia de la información, un síntoma inequívoco de «infoxicación» (intoxicación de información).

Por esto, es importante conocer las herramientas necesarias para navegar por este mar de información, aprender a evaluar fuentes, reconocer sesgos y ser críticos ante la información que recibimos diariamente. La alfabetización digital y el pensamiento crítico son nuestros mejores aliados para enfrentar los desafíos de la era digital y **tomar decisiones informadas en un mundo cada vez más complejo.**

De las «*fake news*» a los trastornos de la información

Aunque el término «*fake news*» es ampliamente utilizado para describir **«noticias falsas» o contenido digital manipulado** que circula principalmente en internet, actualmente diversas fuentes desaconsejan su uso.

La Unión Europea, las Naciones Unidas y la Red Internacional de Verificadores (IFCN) sugieren evitar este término porque es vago, no describe fielmente el alcance del fenómeno, ni los escenarios en los cuales se plantea.

conceptos básicos

«ES UN TÉRMINO VAGO Y AMBIGUO QUE ABARCA TODO, DESDE FALSO EQUILIBRIO (NOTICIAS REALES QUE NO MERECEN NUESTRA ATENCIÓN), PROPAGANDA (DISCURSO ARMADO DISEÑADO PARA APOYAR A UN PARTIDO SOBRE OTRO) Y DESINFORMACIÓN (INFORMACIÓN DISEÑADA PARA SEMBRAR DUDAS Y AUMENTAR LA DESCONFIANZA EN LAS INSTITUCIONES)».

ETHAN ZUCKERMAN

Y, mucho más importante que esto, es el hecho de que algunos políticos alrededor del mundo han utilizado el término «*fake news*» para describir contenidos e incluso a medios de comunicación cuya cobertura no les es favorable.

¿Qué podemos hacer?

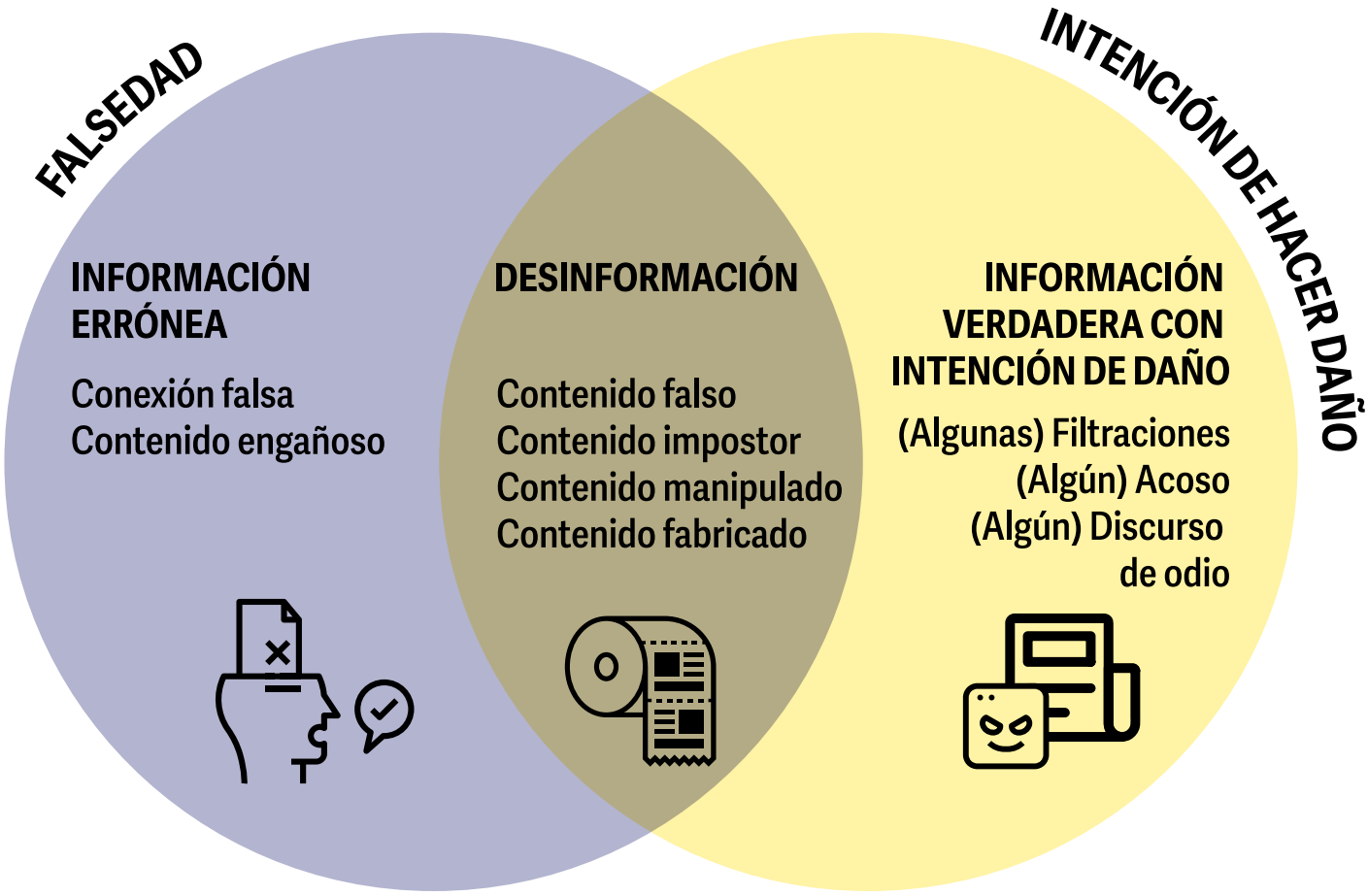
«DE ESTA MANERA, [EL TÉRMINO «*FAKE NEWS*»] SE ESTÁ CONVIRTIENDO EN UN MECANISMO POR EL CUAL LOS PODEROSOS PUEDEN REPRIMIR, RESTRINGIR, SOCAVAR Y ELUDIR LA PRENSA LIBRE».

INFORME DGI(2017)09, CONSEJO DE EUROPA

conceptos básicos

Aquí es donde entra en juego un nuevo marco conceptual que es necesario conocer, comprender y difundir: el de los trastornos de la información.

Información errónea *(misinformation)*: **contenido**

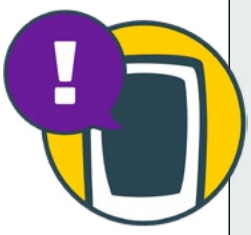


FUENTE: CAZADORES DE *FAKE NEWS*

conceptos básicos

1

incorrecto, pero que se produce o distribuye espontáneamente, **sin intención de dañar.** Puede incluir errores en notas periodísticas o la distribución de contenido falso por creer que es verdadero.



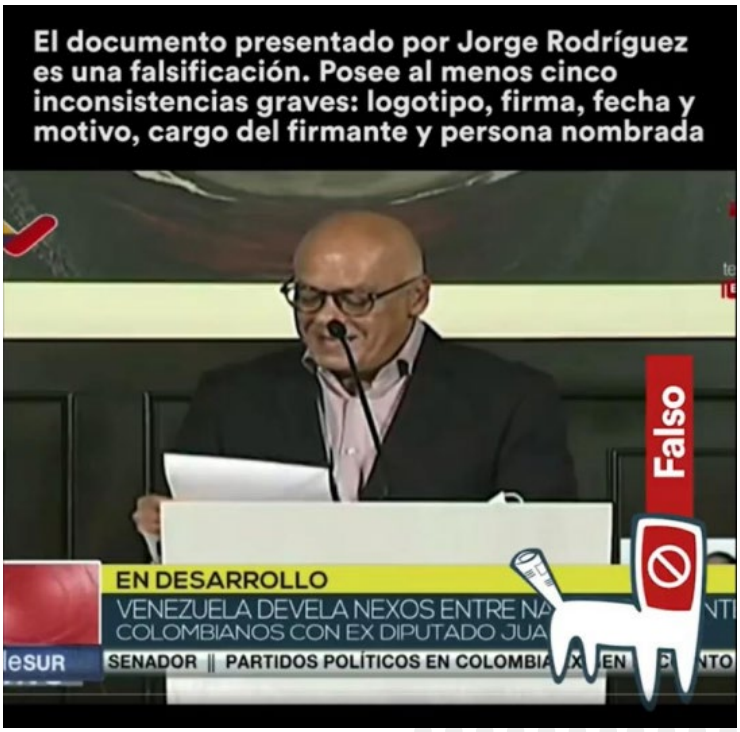
Caso: *Esta fotografía ha sido compartida en redes sociales y servicios de mensajería instantánea, junto con la información de que es un hombre que quedó sin memoria después de un atraco y está recluido en el Hospital Razetti.*

Aplicando técnicas de OSINT pudimos detectar que esa misma fotografía había sido compartida desde hace años en Honduras, Nicaragua, Perú y México, además de Venezuela, a pesar de que no hay datos que confirmen que el hombre realmente fue encontrado en esas condiciones en alguno de esos países. Este es un caso clásico de información errónea que, aunque se difunde sin intención de generar daño —pues generalmente es

conceptos básicos

compartido como un acto de solidaridad y ayuda a la supuesta víctima—, no deja de ser una mentira.

2 **Desinformación** *(disinformation)*: información **completamente falsa, engañosa, descontextualizada o falsificada** y que, simultáneamente, **fue creada con el objetivo de causarle daño a un objetivo.** Se caracteriza por ser **diseñada intencionalmente apelando a sesgos y emociones**, y a menudo se viraliza por intereses políticos, ideológicos o económicos.



Jorge Rodriguez presentó un documento como prueba que Juan Guaidó nombró a Biagio Garófalo como Coordinador Municipal VP en Anaco

Luego de analizar el documento presentado, en Cazadores de *fake news* concluimos que se trató de una falsificación que tenía inconsistencias graves en el logotipo, firma, cargo, fecha y motivo, por lo que fue **creado y expuesto para atacar a un adversario** político.

conceptos básicos

3 **Información verdadera con intención de daño** *(malinformation)*: **contenido real** que se difunde de manera maliciosa **con la intención de causar daño** a un objetivo específico. Un ejemplo clásico es la divulgación de detalles verdaderos sobre corrupción, o información escandalosa sobre un candidato justo antes de un proceso electoral, con el fin de dañar su imagen pública.

Caso: *Once días antes de las elecciones de 2016 en Estados Unidos, el FBI reabrió la investigación sobre una filtración de correos electrónicos proveniente de un servidor privado de Hillary Clinton, donde almacenó correos oficiales durante su mandato como Secretaria de Estado.*

La investigación se llevó a cabo con base en un escándalo sobre la filtración real de información sensible, que hicieron cuestionar el manejo de información por parte de Clinton y su equipo. Sin embargo, la filtración también tuvo como intención perjudicar su campaña electoral y pudo atribuirse a «Fancy Bear», una unidad de hackers que, según investigadores, ha recopilado inteligencia en nombre del gobierno ruso. Este es un caso emblemático de interferencia política externa y divulgación de información verdadera con intención de daño.

conceptos básicos

Más allá del mensaje: develando los trastornos de información

Cuando evaluamos los trastornos de información, es crucial ampliar nuestra perspectiva más allá del mensaje en sí. Debemos explorar el origen del mensaje, los canales de transmisión y, sobre todo, desentrañar **las verdaderas intenciones detrás de su difusión.**

El actor que emite un mensaje es tan relevante como el propio contenido, ya que brinda pistas sobre las posibles intenciones ocultas y ayuda a discernir si nos enfrentamos a un caso de desinformación.

Solo al tener una visión completa del panorama podemos determinar si nos encontramos frente a **información errónea, información verdadera con intención de daño o desinformación.**

conceptos básicos

Las formas de los trastornos de la información

Bajo el paraguas conceptual de los trastornos de la información, podemos identificar diferentes **tipos de contenido que intentan manipular nuestra percepción** sobre algún tema. Comprender estos conceptos permite detectar rápidamente cuándo estamos expuestos a alguna forma de trastorno informativo.

Faktabaari define a la propaganda como «una forma amplia de influencia que busca persuadir a una audiencia para que actúe de acuerdo con los objetivos del propagandista». Se basa en la transmisión de narrativas simplistas, generalmente emocionales, cuyo sello distintivo es la **manipulación psicológica**.

Las teorías conspirativas son explicaciones de eventos o situaciones que postulan la existencia de una **conspiración orquestada por grupos poderosos** con motivaciones a menudo políticas. Quienes las impulsan suelen contradecir el consenso de expertos y reforzarlas mediante más narrativas que usualmente no están basadas en evidencias.

«DETECCIÓN DE TEORÍAS CONSPIRATORIAS». COMISIÓN EUROPEA

1	2	3	4	5	6
UNA SUPUESTA TRAMA SECRETA	UN GRUPO DE CONSPIRADORES	«PRUEBAS» QUE PARECEN APOYAR LA TEORÍA DE LA CONSPIRACIÓN	SUGIEREN FALSAMENTE QUE NADA ES ACCIDENTAL Y QUE LAS COINCIDENCIAS NO EXISTEN	DIVIDEN EL MUNDO ENTRE BUENOS Y MALOS	UTILIZAN A DETERMINADAS PERSONAS Y GRUPOS COMO CHIVOS EXPIATORIOS

La **manipulación digital** es la modificación o alteración de imágenes o videos mediante programas de computadora. Recientemente, la **Inteligencia Artificial (IA) generativa** permite crear imágenes ficticias desde cero, sin tener que manipular contenido preexistente. Esto hace que ciertas imágenes

sintéticas parezcan reales, dificultando su detección.

En el área tecnológica las redes sociales también son un entorno susceptible a la **manipulación de la plataforma**, en la que actores desinformantes ejecutan tácticas individuales o en grupo para posicionar mensajes o dar una falsa relevancia a los temas de su interés.

No podemos olvidarnos de las **mentiras y falsedades.** Estas son **informaciones incorrectas** o no verificadas que a veces se difunden espontáneamente, sin intención de hacer daño y, en otras ocasiones, deliberadamente. En el primer caso, estamos frente a «información errónea», mientras que en el segundo caso, nos encontramos ante un proceso de desinformación diseñado para tergiversar la realidad con un objetivo específico.

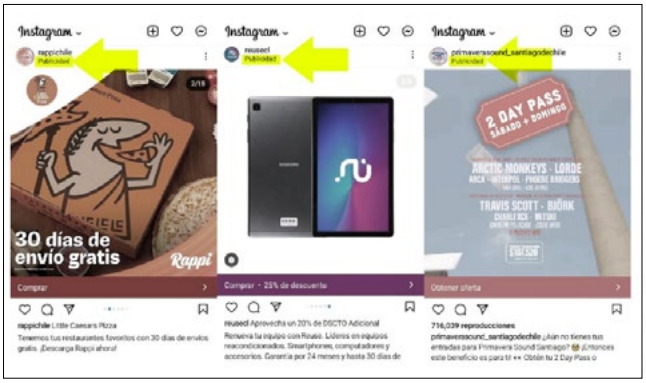
Los **rumores** son informaciones poco específicas que **no han sido o no pueden ser investigadas para verificar su veracidad**. Pueden circular espontáneamente o ser fabricadas para desinformar. Suelen responder a una necesidad pública de alivio, explicación, o justificación y vincularse a informes o noticias en desarrollo que aún no han sido verificadas como verdaderas.

Finalmente, tenemos el **contenido hiperpartidista.** Estos son mensajes con sesgos de origen que se transmiten a través de portales y plataformas relacionadas con partidos políticos, con el objetivo de **persuadir políticamente** a quienes los reciben.

mensajes que pueden confundir

Mensajes que pueden confundir

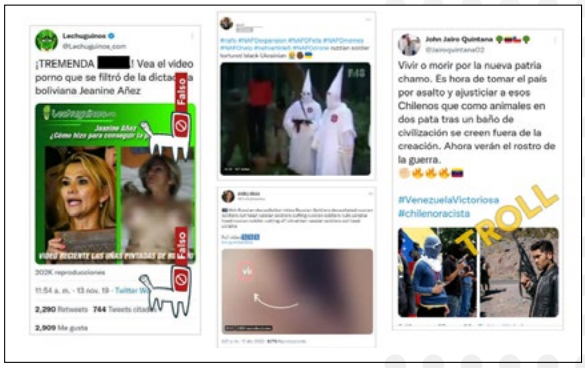
1 **PUBLICIDAD ABIERTA**



2 **PROPAGANDA ENCUBIERTA**

3 **MEDIOS FALSOS**

4 **CONTENIDO DE ODIO**



5 **CLICKBAITS**

conceptos básicos

Descubriendo el mundo de OSINT: inteligencia de fuentes abiertas

OSINT son las siglas en inglés de **Open Source INTelligence** (Inteligencia de Fuentes Abiertas) y se refiere a un conjunto de técnicas y herramientas en línea, que son de acceso abierto, que pueden ser usadas para **recopilar información pública, analizar datos y convertirlos en conocimiento útil.**

Las herramientas OSINT permiten buscar fotografías, mapas, rastros digitales, búsquedas más específicas, rastrear buques o aviones, identificar zonas afectadas por incendios y muchas cosas más. Su uso abarca el ámbito social, político, militar, marketing, etc.

PUEDES HACER OSINT EN MOTORES DE BÚSQUEDA COMO GOOGLE, YANDEX, BING Y TINEYE: Y TAMBIÉN EN REDES SOCIALES COMO TWITTER, FACEBOOK, INSTAGRAM, YOUTUBE, LINKEDIN, VKONTAKTE, TIKTOK Y WEIBO.

Dónde se aplica

La Verificación con Fuentes Abiertas (OSINT) sirve para realizar investigaciones de una manera eficaz, permitiendo rastrear el origen de una información o conseguir detalles sobre una empresa, persona física o cualquier hecho que se desee investigar. Por esto es útil en:

- Lucha contra la desinformación
- Lucha contra cibercrimen
- Lucha contra el terrorismo
- Ciberseguridad y ciberdefensa
- Análisis de opinión pública
- Marketing digital

Servicios que pueden usarse

Los servicios que se pueden usar para hacer OSINT son muy variados. Por un lado se pueden aplicar estrategias de OSINT en los motores de búsqueda como Google, Yandex, Bing, TinEye. Esta búsqueda puede ser sencilla o con operadores y también puede ser una búsqueda inversa de imágenes.

Otras plataformas para hacer OSINT son las redes sociales. Se pueden realizar búsquedas sencillas y algunas con operadores en redes como X/Twitter, Facebook, Instagram, YouTube, LinkedIn, VKontakte, TikTok y Weibo.

También las apps de mensajería instantánea pueden servir para hacer OSINT, principalmente Telegram. WhatsApp, por el contrario, es un servicio mucho más cerrado y es más difícil usar técnicas OSINT en ella, aunque es posible realizar rastreo de información de interés en sus canales y grupos públicos.

¿Qué métodos y herramientas podemos usar para investigar con OSINT?

Existen distintas herramientas que se pueden emplear para hacer OSINT, algunas gratis, otras pagas. Sin embargo, hay tres métodos básicos y comunes que deben ser considerados al iniciar una verificación con fuentes abiertas:

Operadores de búsqueda: son símbolos y comandos específicos que permiten refinar los resultados que obtienes al realizar búsquedas en línea. Actúan como filtros avanzados que, a diferencia de las opciones estándar en la barra de búsqueda, deben ser introducidos manualmente dentro de la consulta. Este enfoque personalizado facilita la localización más precisa de la información que estás buscando.

Geolocalización: es una metodología usada para hallar el sitio específico en el que fue tomada una fotografía o un video, usando combinaciones de técnicas de búsqueda y servicios de mapas en línea como Google Earth, Yandex Maps o Mapillary.

Búsqueda de rastros en redes sociales: Que permiten hallar la publicación exacta de un tuit o nombres de usuario previos de una cuenta, entre otras cosas.

Google

Yandex

Bing

TinEye

Motores de Búsqueda
Búsqueda (sencilla o con operadores)
Búsqueda inversa de imágenes

Facebook Instagram X YouTube

LinkedIn Vkontakte TikTok Weibo

Redes Sociales
Búsqueda (sencilla, algunas con operadores)

Whatsapp Telegram Messenger Meet

Mensajería instantánea

EN LA APP CAZADORES MANTENEMOS UN LISTADO ACTUALIZADO DE HERRAMIENTAS Y SERVICIOS DISPONIBLES PARA OSINT.

ESCANEA EL QR
O HAZ CLICK

conceptos básicos

Metodología de una investigación OSINT

Hay cinco pasos que se deben seguir al momento de plantearse ejecutar una investigación OSINT: dirección, recolección, procesamiento, análisis y difusión.

1. **Dirección:** Determinar qué es lo que se quiere buscar y cuáles son los objetivos que persigue la investigación. En esta etapa se hace una toma de datos iniciales para ver hacia dónde se puede dirigir la información.
2. **Recolección:** Una búsqueda más avanzada de datos e información empleando las herramientas disponibles de OSINT.
3. **Procesamiento:** La información que se consigue se procesa y se prepara para su análisis. En esta parte de la investigación se pueden elaborar tablas, gráficos o mapas para tener una mejor visualización de los datos obtenidos.
4. **Análisis:** Los datos finalmente se evalúan y analizan para generar conocimiento sobre el tema y se genera el informe con los objetivos planteados.
5. **Difusión:** En esta parte se entrega la información y las conclusiones de la investigación.

conceptos básicos

Cómo saber si una información es confiable

Organizaciones anti-desinformación ucranianas comenzaron a difundir a principios de 2021 la imagen de la «Flor Antifake», una guía de 7 preguntas que pueden ser usadas por cualquier persona para **evaluar contenido potencialmente desinformativo**. Es una forma sencilla de fomentar el **pensamiento crítico** en la sociedad civil.

SEGUNDA PARTE

Aprende a hacer búsquedas como un experto

02

02

SEGUNDA PARTE

Aprende a hacer búsquedas como un experto

EL GOOGLE DORKING, TAMBIÉN LLAMADO GOOGLE HACKING, ES UNA TÉCNICA QUE UTILIZA CONSULTAS DE BÚSQUEDA AVANZADAS PARA ACCEDER A INFORMACIÓN OCULTA EN GOOGLE.

AVG.COM

ESCANEA EL QR O HAZ CLICK PARA ACCEDER A LA APP CAZADORES

En la era digital, los motores de búsqueda se han convertido en una herramienta cotidiana que guía nuestro camino en línea. Desde respuestas rápidas hasta profundas investigaciones, cualquier ser humano confía en la capacidad de los buscadores para encontrar información relevante y los investigadores y periodistas no son la excepción. Los motores de búsqueda son minas de oro de datos valiosos que, en este capítulo, te enseñaremos a aprovechar.

¿Alguna vez has oído hablar de los **«Google Dorks»**? Estas son palabras claves que funcionan como filtros para ubicar información específica en Internet, como una llave maestra que abre todas las puertas, incluso las que no están señaladas. Utilizar operadores y comandos en tus búsquedas puede transformar resultados genéricos en hallazgos especializados. Los periodistas e investigadores deben aprovechar esta función para desenterrar esa información que otros pasan por alto.

Operadores: más que palabras clave

Si bien los operadores pueden parecer intimidantes al principio, son recursos esenciales en la caja de herramientas de investigación OSINT y no se necesita ser un programador para usarlos. Añadir comillas, guiones o términos como `«site:», «inurl:» o`

hacer búsquedas como un experto

«`filetype:`» a una consulta puede hacer la diferencia entre un sinfín de información inútil y exactamente lo que estás buscando. Así, los resultados son precisos y enfocados, lo que ahorra tiempo y te guía a los datos importantes de manera eficiente.

Para incluir esta herramienta en la rutina de investigación con facilidad, en la sección «**Operadores de búsqueda**» de la App Cazadores mantenemos una lista actualizada con los Dorks más comunes y útiles para la investigación OSINT. Allí también están disponibles los motores de búsqueda en los que funcionan y otras plataformas donde los puedes utilizar.

En un mundo donde la desinformación y las noticias falsas se propagan rápidamente, las búsquedas precisas son fundamentales para los periodistas y los investigadores porque nos permiten filtrar fuentes y desentrañar la verdad detrás de las narrativas distorsionadas. **La habilidad de rastrear fuentes y verificar información se vuelve más poderosa que nunca** cuando dominamos las técnicas de búsqueda avanzada.

Investigación OSINT de textos

En la era digital, donde la información fluye a velocidades vertiginosas, la investigación de textos se ha convertido en una habilidad fundamental para discernir entre los hechos y la desinformación. En este contexto, los operadores de búsqueda pueden ser usados para filtrar contenidos específicos y acelerar el hallazgo de contenido difícil de acceder, mejorando la precisión y eficiencia en las búsquedas.

hacer búsquedas como un experto

1 Paso 1: Identificar el tipo de texto

El primer paso en la investigación de textos con OSINT es determinar el tipo de contenido que estás analizando. ¿Se trata de una simple cadena de texto, un fragmento de conversación de WhatsApp, una captura de pantalla de un post en redes sociales o un extracto de un artículo? Cada tipo de texto puede requerir un enfoque específico para obtener resultados relevantes.

2 Paso 2: Potenciando la búsqueda con operadores avanzados

Los operadores avanzados como el guion (-), AND, comillas y las fechas (after: before: en Google, o since: until: en Twitter) son herramientas poderosas para perfeccionar tus búsquedas y obtener resultados más relevantes y precisos. Por ejemplo, el guion (-) te permite excluir palabras no deseadas de tus resultados, mientras que el operador AND combina términos para encontrar resultados que contengan ambas palabras. Las comillas ("") son útiles para buscar frases exactas, lo que es esencial para verificar afirmaciones. Las fechas son especialmente valiosas para investigar eventos en un rango de tiempo específico.

3 Paso 3: Búsqueda con palabras clave

Para investigar a fondo un texto, es crucial identificar sus principales palabras o frases clave. Son los términos específicos que representan la esencia de la información que estás buscando. Al utilizar operadores como el guion (-), el operador AND y las comillas ("") en una búsqueda con términos clave, puedes refinar los resultados para encontrar información que se ajuste con precisión a tu intención de búsqueda.

hacer búsquedas como un experto

4 Paso 4: Uso de operadores en fuentes oficiales

Una vez que hayas definido el tipo de texto y verificado su legitimidad como fuente oficial, los operadores de búsqueda pueden ser tus aliados. Operadores como «`site:`» o «`inurl:`» te permiten limitar tus búsquedas a un sitio web o alguna cuenta oficial en redes sociales, lo que resulta especialmente útil al analizar información proveniente de fuentes oficiales.

La investigación de textos con OSINT se ha vuelto **una habilidad indispensable en la era de la información**. Los operadores de búsqueda son la clave para filtrar y obtener información precisa y relevante en medio del mar de datos en línea. Al identificar el tipo de texto, usar operadores en textos oficiales, buscar palabras clave y emplear operadores avanzados, los investigadores y periodistas pueden optimizar sus resultados y distinguir los hechos entre tanta información.

Artículos de opinión y algunos análisis no basados en datos

Una capacidad imprescindible para evaluar información es comprender las diferencias entre artículos de opinión e información basada en hechos. En el artículo de opinión, el autor comparte sus perspectivas personales o profesionales sobre un tema en particular. Estos textos están intrínsecamente cargados de elementos subjetivos, ya que reflejan una **interpretación individual de los acontecimientos**. Por ejemplo, un artículo que argumenta a favor o en contra de una política gubernamental puede estar basado en la percepción y la experiencia única del autor, lo que hace que este tipo de contenido no siempre sea verificable o

hacer búsquedas como un experto

desmentible en términos estrictamente factuales.

Por otro lado, la información basada en hechos se caracteriza por su **objetividad y verificabilidad.** Estos textos se apoyan en evidencia concreta, datos verificables y testimonios confiables. Si consideramos un artículo que informa sobre los resultados de un estudio científico respaldado por pruebas y datos precisos, estamos frente a información que puede ser corroborada por otros investigadores y expertos en el campo. En contraste, los artículos de opinión no pueden ser desmentidos o verificados de la misma manera, ya que están arraigados en la interpretación de un individuo. Reconocer esta distinción es fundamental para consumir información de manera informada y crítica en la era digital.

Denuncias

Una denuncia es una comunicación formal en la que una persona o entidad informa sobre un acto ilegal, inmoral o irregular a las autoridades pertinentes o al público en general. Este acto busca poner en conocimiento hechos que pueden ser perjudiciales, peligrosos o contrarios a las normas establecidas. En el ámbito periodístico, las denuncias son herramientas cruciales para exponer la corrupción, abusos de poder y otras injusticias, lo que puede contribuir a una sociedad más justa y responsable.

> **UNA DENUNCIA NO SIEMPRE ES SINÓNIMO DE VERDAD ABSOLUTA, Y EXISTE LA POSIBILIDAD DE QUE ALGUNAS SEAN FALSAS O ESTÉN MOTIVADAS POR INTERESES OCULTOS.**

Sin embargo, los periodistas tienen la **responsabilidad de verificar la confiabilidad** tanto de la fuente de la denuncia como de los datos proporcionados antes de publicar la información. Sobre todo, en el contexto de denuncias que no se presentan ante los organismos

hacer búsquedas como un experto

competentes, sino que se hacen públicas a través de las cuentas en redes sociales de individuos comunes.

La verificación cuidadosa de la fuente y de los datos es esencial para garantizar la integridad y credibilidad de la información presentada. Además, publicar información no verificada o falsa puede tener graves consecuencias legales y reputacionales, tanto para los periodistas como para las partes involucradas. Por lo tanto, la **verificación rigurosa es una práctica esencial para mantener la ética periodística** y brindar a la audiencia información precisa y confiable.

Filtraciones

Las filtraciones de datos se refieren a la **divulgación no autorizada de información confidencial o privada**, que puede incluir correos electrónicos, documentos internos, conversaciones de mensajería, entre otros. Estas filtraciones a menudo tienen el propósito de exponer irregularidades, abusos o corrupción, y pueden tener un impacto significativo en la sociedad y en las organizaciones afectadas. Sin embargo, la veracidad y autenticidad de dichos datos deben ser rigurosamente examinadas antes de considerar que se trata de información confiable.

Para verificar las filtraciones de datos, es crucial evaluar tanto el origen de los datos como su contenido. Si se tiene acceso a la base de datos original, es importante **comprobar la legitimidad de la fuente** y asegurarse de que los datos no hayan sido alterados o manipulados. Si se trata de capturas de pantalla o imágenes de documentos, es necesario verificar la autenticidad de la imagen, ya que éstas pueden ser editadas o modificadas para **distorsionar la información**.

hacer búsquedas como un experto

Al verificar la autenticidad de las filtraciones, se pueden considerar varios pasos, como examinar los metadatos de los archivos para determinar si han sido modificados, comparar la información con otras fuentes fiables o buscar inconsistencias en los datos presentados. Además, se puede buscar la opinión de expertos en ciberseguridad o forenses digitales para evaluar su validez.

Si bien las filtraciones de datos pueden ser valiosas para exponer la verdad y la transparencia, es fundamental llevar a cabo una verificación exhaustiva para garantizar la autenticidad y veracidad de los datos. Esto implica evaluar tanto la fuente como el contenido de los datos y **considerar la posibilidad de manipulación de imágenes**. La integridad y credibilidad de la información dependen de una rigurosa verificación de las fuentes y los datos antes de utilizarlos en la toma de decisiones o en divulgarlos como información verdadera.

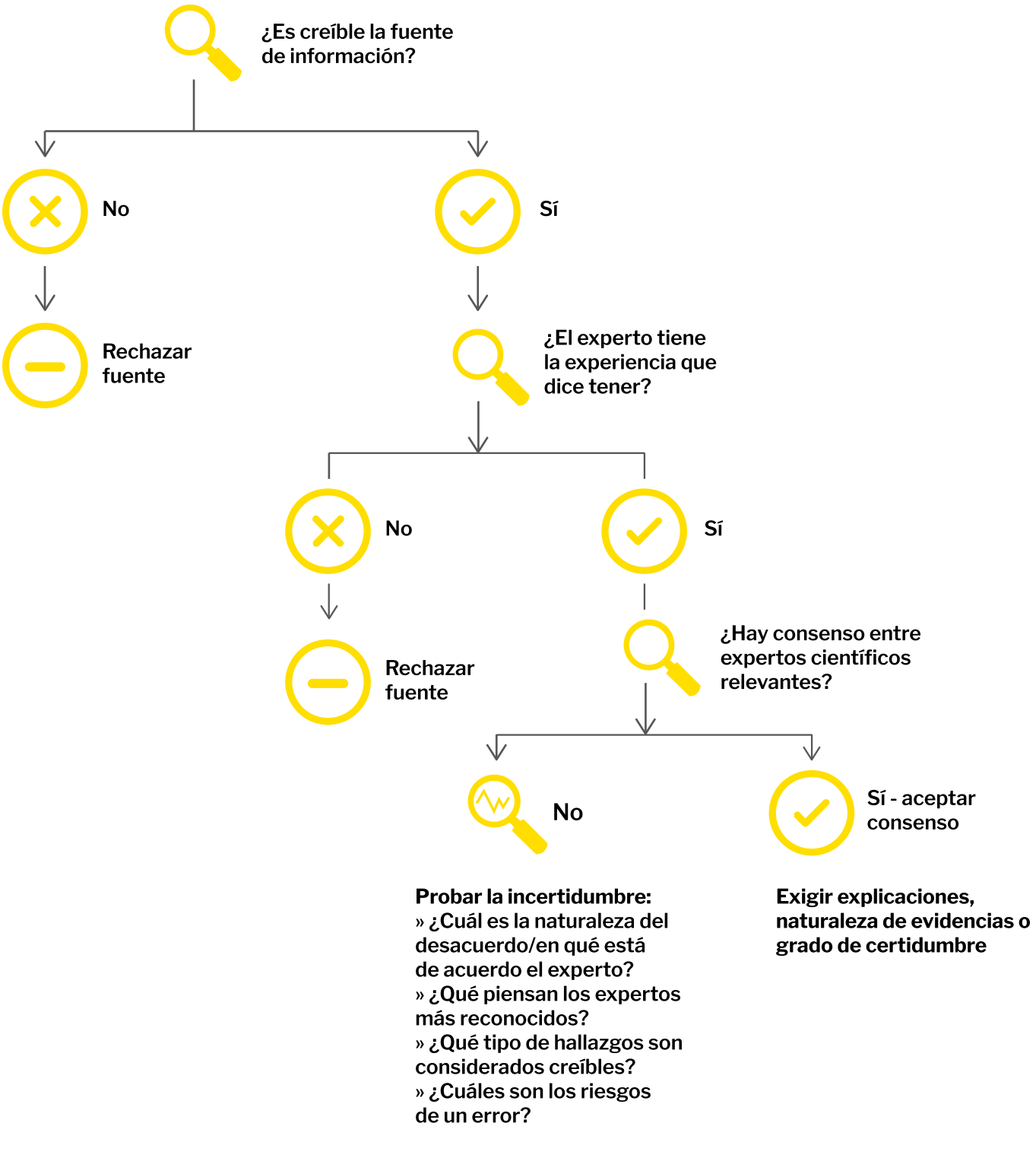
Información científica

La pandemia del COVID-19 planteó un nuevo reto a las iniciativas de verificación de todo el mundo. Si bien la comunidad científica internacional se alineó en la búsqueda urgente de una vacuna que permitiera hacerle frente a la enfermedad, una minoría de fuentes –en algunos casos voceros de la salud o científicos– amplificaron teorías de conspiración tanto sobre la enfermedad como de su tratamiento, que con el tiempo se demostraron erróneas.

Es por ello que algunas iniciativas plantearon la necesidad de estipular árboles de decisión similares al presentado en la siguiente página para evaluar información científica, especialmente en situaciones de emergencia:

hacer búsquedas como un experto

ÁRBOL DE DECISIÓN PARA EVALUAR LA INFORMACIÓN CIENTÍFICA



FUENTE: DIGITAL INFORMATION LITERACY GUIDE | FAKTABAARI EDU. 2022

hacer búsquedas como un experto

Investigación OSINT de audios

La investigación de contenido en formato de audio presenta **desafíos únicos para los investigadores OSINT**. Aunque la tecnología avanza rápidamente, los buscadores aún no permiten hacer «búsquedas inversas» de clips de sonido, lo que complica la tarea de localizar información específica sobre audios. Además, la creciente capacidad de la inteligencia artificial para generar audios falsos con la voz de cualquier persona agrega un nivel adicional de complejidad y desconfianza en la autenticidad de los contenidos.

Para abordar la verificación de audios, los investigadores pueden optar por investigar estos contenidos **tal como se investigan textos virales**. Para ello, es conveniente identificar palabras clave y nombres mencionados en el audio y utilizarlos en búsquedas con operadores. Así, se puede obtener información relevante relacionada con los temas o personas discutidas en el registro.

Cuando el audio está en un idioma diferente o se sospeche que el original es en otro idioma, es posible **identificar y traducir palabras clave**, para luego repetir las búsquedas como texto. Esto amplía la posibilidad de obtener resultados relevantes para investigar el origen del audio y poder determinar si está descontextualizado.

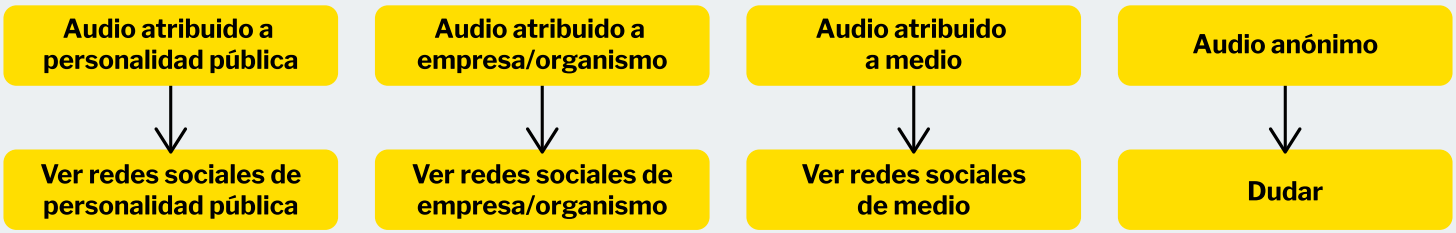
Para información sobre canciones, se pueden hacer búsquedas avanzadas con operadores en conjunto con frases cortas y entrecomilladas de la canción, pudiendo en algunas ocasiones conocer el artista y el título de la canción. Para una identificación más precisa, herramientas como la opción de **audios de Google o aplicaciones como**

hacer búsquedas como un experto

Shazam permiten reconocer canciones, versiones y autores a partir de la reproducción del audio.

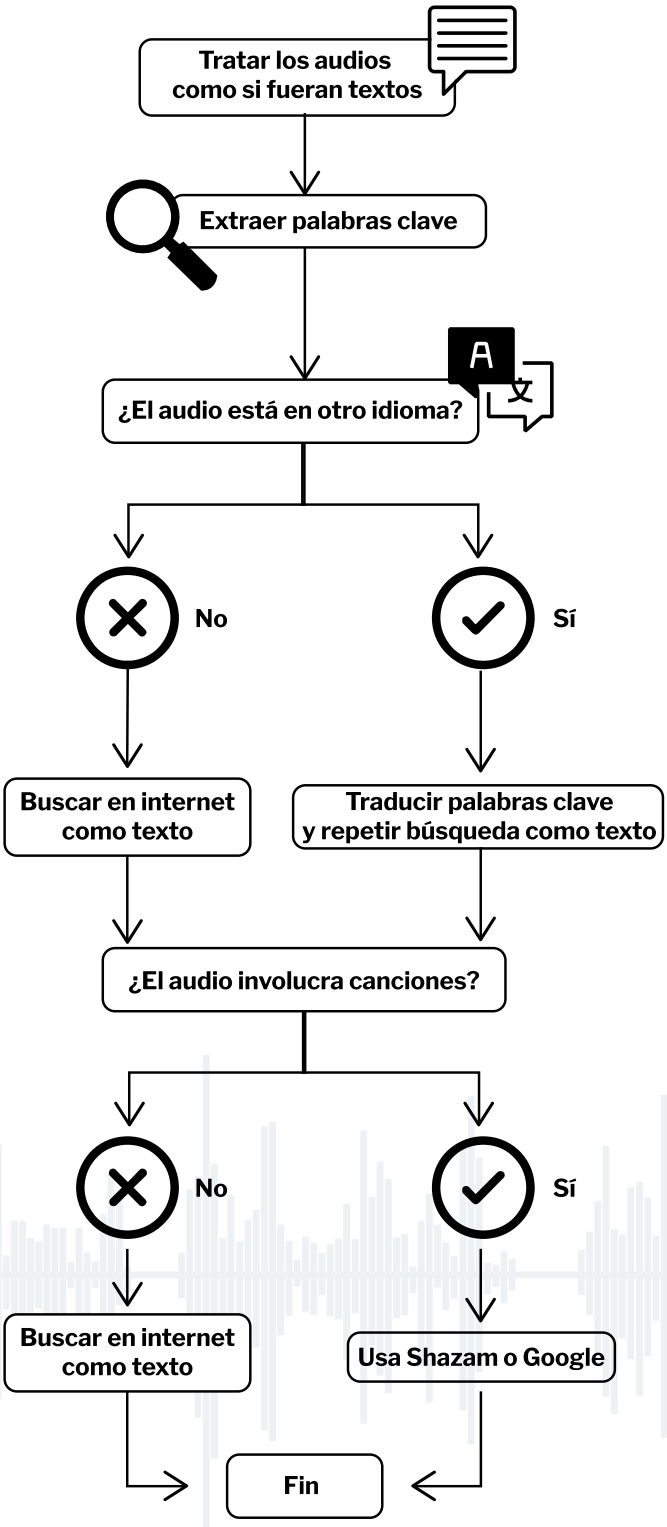
La versatilidad del uso de operadores de búsqueda junto a extractos cortos y entrecomillados de un segmento de audio, no solamente sirve para hallar información sobre canciones. Usando la misma técnica, en ocasiones es posible identificar el origen de audios o videos de discursos famosos de figuras públicas, especialmente cuando fueron escritos y publicados en medios de comunicación, páginas web o blogs.

Cómo determinar la confiabilidad de un audio



hacer búsquedas como un experto

¿Qué pasos seguir para verificar un audio viral?



hacer búsquedas como un experto

EL ÉXITO DE UNA CAMPAÑA DE RUMORES VARÍA EN RELACIÓN CON LA IMPORTANCIA Y AMBIGÜEDAD DE LA SITUACIÓN. MIENTRAS MAYOR SEA LA IMPORTANCIA O LA AMBIGÜEDAD, MAYOR SERÁ EL EFECTO QUE TENDRÁ EN EL COMPORTAMIENTO Y EN LA ACTITUD.

MANUAL DE OPERACIONES PSICOLÓGICAS. EJÉRCITO VENEZOLANO CARACAS, 2006

La dificultad de investigar audios anónimos

La verificación de audios anónimos conlleva un desafío esencial: la falta de credibilidad en la fuente. Frecuentemente, estos audios contienen información falsa, engañosa o una combinación de **datos reales con opiniones o apreciaciones subjetivas de un emisor desconocido.**

A diferencia de la desinformación que circula en forma de textos e imágenes, la desinformación en audio puede darle al receptor una falsa ilusión de cercanía con el autor del registro, o hacerle creer que ha conseguido tener acceso a información privilegiada, en vez de a un engaño. Esta mezcla puede distorsionar la realidad y crear una narrativa engañosa. Por esta razón, los rumores en audio suelen formar parte esencial en **campañas de rumores y operaciones de influencia**, siendo herederos naturales en el mundo digital de los rumores que en el pasado se transmitían de boca en boca en una comunidad.

Además, la naturaleza anónima de muchos audios virales dificulta discernir su origen y autenticidad, lo que complica aún más su verificación. Aunque pueden influir en la percepción de los receptores sobre un tema en particular, es importante resaltar la necesidad de que estos registros pasen por rigurosos procesos de verificación, especialmente cuando no pueden ser atribuidos a ninguna fuente confiable. En última instancia, es importante abordar la investigación de audios virales, especialmente los anónimos, con un **enfoque crítico y basado en hechos**, para evitar la propagación de desinformación y rumores.

hacer búsquedas como un experto

Investigación OSINT de imágenes

Investigar imágenes a través de OSINT brinda oportunidades únicas para verificar su autenticidad y contexto. Aunque se pueden usar las mismas técnicas de investigación de texto para hallar el origen de imágenes, es esencial considerar otros aspectos para el rastreo de información viral en formato gráfico:

1. Cuando las imágenes son presentadas como comunicados oficiales en nombre de una institución, compañía o marca, es crucial **revisar las redes sociales y páginas web oficiales** de dicha entidad. Esto permite confirmar o descartar la autenticidad de la imagen, acudiendo directamente a la fuente.

2. En el caso de imágenes que contienen títulos y leyendas, como portadas de revistas o fotos de programas de televisión, se pueden emplear **operadores de búsqueda con palabras clave** relacionadas al tema u objeto en la imagen y a la información que aparece en texto. Esta técnica puede permitir rastrear el origen de la imagen viral y llevarnos a la publicación original o, por el contrario, a desmentidos existentes sobre el caso evaluado.

3. Una técnica OSINT imprescindible es la **búsqueda inversa de imágenes.** Consiste en cargar la imagen a ser investigada en motores de búsqueda inversa de imágenes como Google Images o TinEye, obteniendo como resultado los sitios web en los que la imagen había sido publicada previamente. Esta técnica puede ayudar a descubrir el origen de la imagen, su autenticidad y si ha sido modificada o sacada de contexto.

Las anteriores son algunas técnicas usadas por investigadores OSINT para rastrear el origen de

hacer búsquedas como un experto

imágenes y distinguir entre auténticas y manipuladas.

La investigación de imágenes mediante **OSINT ofrece varias vías de verificación**. Las fuentes oficiales de instituciones, operadores de búsqueda y la búsqueda inversa de imágenes son esenciales para confirmar la autenticidad y contexto contenido viral atribuido a un actor en específico. La búsqueda inversa de imágenes es particularmente valiosa, ya que no solo permite demostrar que alguna imagen está descontextualizada o es una manipulación digital, sino que también puede servir para rastrear la primera aparición del contenido en Internet, proporcionando información sobre la imagen, tal como apareció originalmente en línea.

Aunque los principales buscadores como Google, Yandex y Bing ofrecen servicios de búsqueda inversa de imágenes, también existen otras herramientas menos conocidas. Los enlaces a muchas de ellas se encuentran listados en la sección «Herramientas» de la App Cazadores junto a una gran variedad de herramientas que regularmente son usadas por investigadores OSINT alrededor del mundo, la mayoría de acceso gratuito.

ESCANEA EL QR O HAZ CLICK PARA ACCEDER A LA APP CAZADORES

Investigación OSINT de videos

La investigación OSINT de videos virales se basa en una combinación de técnicas de investigación de textos (incluyendo el uso de operadores de búsqueda) y búsqueda inversa de imágenes estáticas (capturas de pantalla) tomadas del video que se está investigando. A continuación, se describen algunas técnicas útiles para la investigación OSINT de videos virales:

hacer búsquedas como un experto

1. Presta atención a los **detalles que aparecen durante el video:** Observar detenidamente el entorno del video puede proporcionar pistas valiosas sobre su ubicación o contexto. Se pueden identificar carteles, anuncios, marcas de agua, señales de tránsito, logotipos u objetos que puedan dar pistas sobre dónde fue grabado el video, lo que puede permitir rastrear su origen y determinar su autenticidad.

2. Búsqueda con **palabras clave**: Las palabras clave mencionadas en el video o que resumen su contenido son fundamentales. Realizar búsquedas en línea con estas palabras y combinándolas con operadores de búsqueda, puede conducir a verificaciones previas o información que brinde contexto sobre el video.

3. **Hacer búsqueda inversa de capturas** de momentos distintivos: Identificar momentos clave en el video, tomar capturas de pantalla y luego realizar búsqueda inversa de esas imágenes, puede servir para rastrear la publicación original del video.

La investigación de videos con OSINT requiere un enfoque cuidadoso y metódico, utilizando tanto las técnicas de investigación tradicionales como las herramientas digitales disponibles. Al combinar estos pasos, los investigadores pueden desentrañar la verdadera naturaleza de los videos en línea y contribuir a la identificación y difusión de información veraz y precisa.

Uso de plugin especializado en video: InVID

El principal reto en la verificación de videos virales radica en la obtención de resultados efectivos durante la búsqueda inversa de **capturas de pantalla extraídas del video** en cuestión. Aunque este procedimiento no es intrínsecamente

hacer búsquedas como un experto

ESCANEA EL QR O HAZ CLICK PARA ACCEDER A LA APP CAZADORES

complejo, puede llegar a ser frustrante para algunos investigadores, ya que requiere tomar múltiples capturas de pantalla, realizar búsquedas inversas y, posiblemente, repetir todo el proceso en caso de no obtener resultados concluyentes.

Para facilitar y simplificar esta metodología, se han desarrollado herramientas especializadas como InVID. Esta aplicación facilita la tarea al permitir la **generación de capturas de pantalla** extraidas de distintos momentos del video investigado y realizar búsquedas inversas a estas capturas, con tan solo un par de clicks.

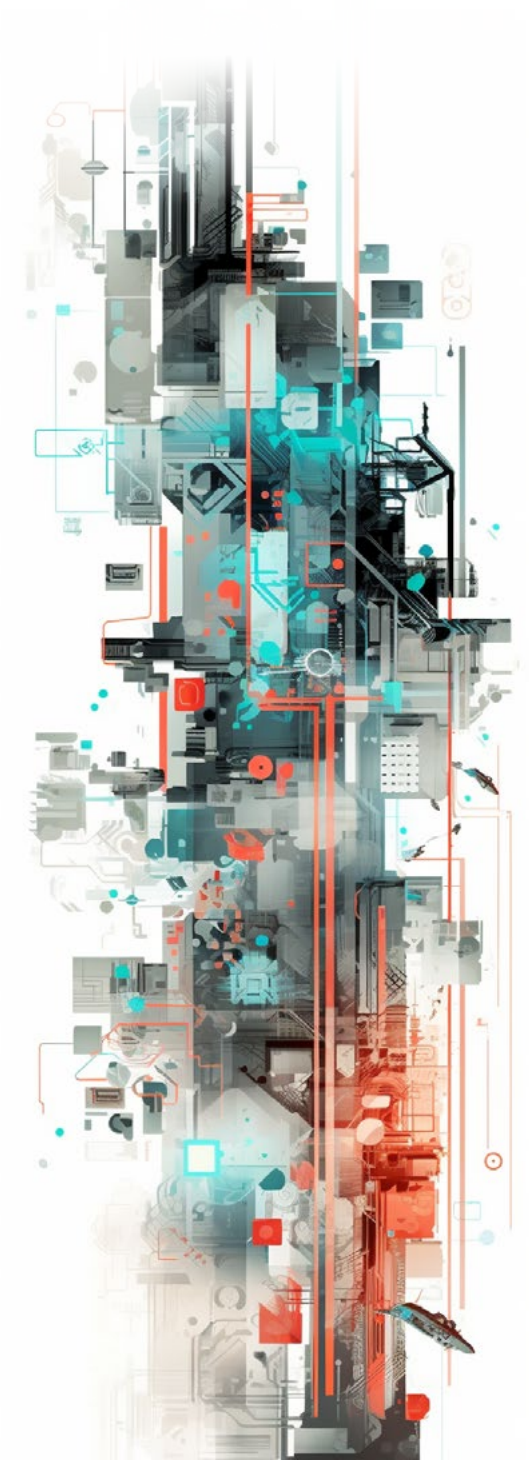
InVID es un proyecto desarrollado por un consorcio de organizaciones, entre las que se encuentran el Centro de Investigación y Tecnología de Hellas (CERTH) - Instituto de Tecnologías de la Información (ITI), Modul Technology GmbH, Universitat de Lleida (UdL) y otras. Es una extensión para navegadores, de acceso gratuito, con la que fácilmente se pueden hacer búsquedas inversas de imágenes, extracción de múltiples capturas de pantalla a videos y listas de metadatos de archivos multimedia, entre otras herramientas.

InVID busca abordar el desafío de verificar la autenticidad y la contextualización de los videos en la web a través de la experiencia y el conocimiento de múltiples organizaciones en tecnologías de la información, análisis de medios y servicios de noticias.

El software permite **verificar la autenticidad y la fuente de los videos, detectar manipulaciones digitales y descontextualizaciones**, así como brindar a los usuarios la capacidad de tomar decisiones informadas sobre la veracidad de los contenidos audiovisuales. Junto con otras herramientas

hacer búsquedas como un experto

disponibles en la App Cazadores, InVID ofrece a los profesionales de la investigación una solución integral para abordar los desafíos de la verificación de videos.



Contenido generativo creado con IA

En 2023, la inteligencia artificial (IA) ha demostrado su capacidad para generar contenido multimedia convincente —como textos, audios, imágenes y videos— a través del uso de «Redes Neuronales (de distintos tipos)», que son **sistemas de aprendizaje automático capaces de crear contenido** que se asemeja al generado por humanos. Aunque esta tecnología tiene aplicaciones beneficiosas, también presenta riesgos en cuanto a la desinformación.

La llegada de la IA ha hecho necesario diferenciar las manipulaciones digitales tradicionales del contenido generativo creado con GAN (Redes Generativas Antagónicas), lo que ha dado lugar a los términos «*shallowfake*» y «*deepfake*». Los *shallowfakes* son **manipulaciones de contenido multimedia que no involucran tecnologías basadas en IA**, GAN ni algoritmos de aprendizaje profundo *(deep learning)*. Son aquellos creados con software de edición de video y generalmente son alteraciones de contenidos existentes.

Por otro lado, los *deepfakes* sí son el **resultado del uso de IA, algoritmos de aprendizaje profundo y GAN**. Las imágenes y videos obtenidos, que pueden simular contenidos reales, generalmente son creados desde cero y a diferencia de los *shallowfakes*, no suelen ser modificaciones de contenido existente.

hacer búsquedas como un experto

Shallowfakes vs. Deepfakes:

Shallowfakes

- Generados mediante edición manual.
- No requieren IA, GAN ni algoritmos de aprendizaje profundo.
- Menos complejos tecnológicamente.
- Alteraciones parciales de contenido multimedia existente.
- Creados con software básico de edición de video.
- Mayor intervención humana en el proceso de manipulación.

Deepfakes

- Utilizan algoritmos de aprendizaje profundo y GAN.
- Requieren IA avanzada.
- Mayor complejidad tecnológica.
- Pueden crear contenido altamente realista y convincente.
- Utilizan técnicas automatizadas.
- Mínima intervención humana en la generación del contenido.

En el combate contra la desinformación, resulta crucial entender las repercusiones de las tecnologías generativas y su capacidad para distorsionar la información que recibimos. Tanto los *shallowfakes* como los *deepfakes* pueden emplearse para desinformar y manipular el discurso en línea, lo cual erosiona la confianza en el contenido que consumimos. Para distinguir entre fotografías y videos auténticos y aquellos que han sido manipulados, es imperativo familiarizarse con algunas características de los diversos tipos de contenido audiovisual alterado.

Softwares más populares para la generación de textos con IA

Los generadores de texto con IA operan automáticamente al basarse en contenido previamente introducido por los usuarios. Estas herramientas no operan con magia, sino a través del **análisis de**

hacer búsquedas como un experto

grandes volúmenes de datos, de los cuales extraen conocimiento y aprenden patrones lingüísticos, lo que les permite generar textos que emulan el lenguaje humano. Su funcionamiento es análogo al proceso que sigue un escritor humano cuando realiza investigaciones previas para fundamentar su obra: los generadores toman información de bases de datos preestablecidas y crean contenidos acordes a las instrucciones proporcionadas por sus usuarios.

Los principales generadores de texto con IA son:

OpenAI

ChatGPT:

Herramienta de OpenAI basada en la arquitectura GPT, capaz de producir contenido coherente y contextual en respuesta a preguntas e instrucciones.

Bard

Google Bard:

IA de Google Research que crea textos en distintos estilos y temas utilizando modelos de lenguaje avanzados y permitiendo guiar la generación con pistas y directrices.

Bing

Bing Chat:

Herramienta de Microsoft que utiliza IA para generar respuestas en tiempo real en chats. Se integra con aplicaciones de mensajería y automatiza respuestas a consultas frecuentes mediante procesamiento del lenguaje natural.

Estos generadores de texto con IA tienen un alto valor en la creación de contenidos y la generación de respuestas en tiempo real. Aunque no sustituyen por completo la creatividad humana, su capacidad para mejorar la eficiencia y la productividad los hace herramientas poderosas, lo que plantea dudas sobre su potencial para contribuir a la desinformación y la manipulación de información.

hacer búsquedas como un experto

Softwares para generar imágenes con IA

Los generadores de imágenes con IA han revolucionado la creación visual a partir de texto, utilizando algoritmos de aprendizaje automático para producir imágenes originales y, a veces, muy realistas. Estos programas combinan estilos, conceptos y atributos para dar vida a imágenes coherentes basadas en descripciones en lenguaje natural. Tres de los generadores de imágenes más destacados son:

Midjourney utiliza IA y redes neuronales para crear imágenes coherentes desde descripciones textuales, enfocándose en la interpretación precisa y la esencia de los conceptos solicitados.

stability.ai usa redes neuronales para convertir descripciones textuales en imágenes realistas, con énfasis en la fidelidad visual y la coherencia, ampliando las capacidades de generación automática.

De los mismos creadores de ChatGPT, **Dall-E** es capaz de generar imágenes a partir de descripciones complejas, empleando redes neuronales avanzadas para explorar ideas abstractas y situaciones, llevando la generación de imágenes a un nivel más profundo y creativo.

Los generadores de imágenes con IA **han transformado la forma en que percibimos y creamos contenido visual**, permitiendo que las palabras cobren vida a través de imágenes coherentes y originales. Sin embargo, como sucede con muchas tecnologías innovadoras, también presentan desafíos y riesgos significativos en el terreno de la desinformación.

Los actores desinformantes pueden aprovechar estas herramientas para **crear imágenes falsas y**

hacer búsquedas como un experto

convincentes que respalden narrativas falsas o engañosas. Esto puede abarcar desde la creación de evidencia falsa para difamar a individuos o empresas, hasta la propagación de información errónea o contenido de odio en redes sociales y sitios web.

Un ejemplo de cómo las imágenes generadas por IA pueden ser utilizadas para la desinformación es el caso en el que **una imagen falsa provocó un *«flash crash»* en la bolsa** de valores. En 2023, una imagen generada con IA y publicada en Twitter mostraba una noticia falsa sobre una explosión en el Pentágono. Esta falsa información se difundió rápidamente y causó una caída momentánea en el mercado de valores.

Este incidente destaca cómo la creación y difusión de imágenes falsas generadas con IA pueden tener un impacto significativo en la percepción pública, la economía y la estabilidad social. Además, la capacidad de estos generadores de imágenes para producir contenido visual

hacer búsquedas como un experto

altamente convincente puede dificultar la detección de imágenes falsas a simple vista.

Mientras que los generadores de imágenes con IA han abierto un mundo de posibilidades creativas y expresivas, también han amplificado los riesgos asociados con la desinformación visual. Abordar estos desafíos requerirá un esfuerzo conjunto para desarrollar técnicas de detección avanzadas, **promover la educación digital y fomentar la responsabilidad en la creación y el consumo de contenido generado por IA.**

Rostros creados con inteligencia artificial

El uso de rostros creados con IA ha encontrado un nuevo nicho en el contexto de la manipulación de las redes sociales, cuando se emplean cuentas falsas generadoras de *spam* o durante operaciones de influencia, convirtiéndose en otra herramienta usada para desinformar y distorsionar la opinión pública.

Una táctica comúnmente empleada para impulsar campañas no auténticas en redes sociales es la utilización de cuentas falsas. El talón de Aquiles de esta estrategia radica en la falta de indicios que sugieran que estas cuentas están siendo operadas por individuos reales. Precisamente para abordar esta debilidad, en varios países de Latinoamérica y el mundo se ha recurrido al uso de **rostros generados con IA**, con el fin de brindar mayor apariencia de legitimidad a redes de cuentas ficticias.

> AL DAR LA APARIENCIA DE PERFILES AUTÉNTICOS, CUENTAS FALSAS QUE HACEN USO DE ROSTROS SINTÉTICOS GENERADOS CON IA PUEDEN ENGAÑAR A LA SOCIEDAD CIVIL, HACIENDO QUE CIERTAS CAMPAÑAS NO AUTÉNTICAS EN REDES SOCIALES PAREZCAN MÁS POPULARES DE LO QUE REALMENTE SON.

hacer búsquedas como un experto

> LA ALFABETIZACIÓN DIGITAL ES IMPORTANTE PARA QUE SOCIEDAD DESARROLLE ESTRATEGIAS DE DETECCIÓN Y CONTRAMEDIDAS EFECTIVAS PARA COMBATIR LA DESINFORMACIÓN EN LÍNEA.

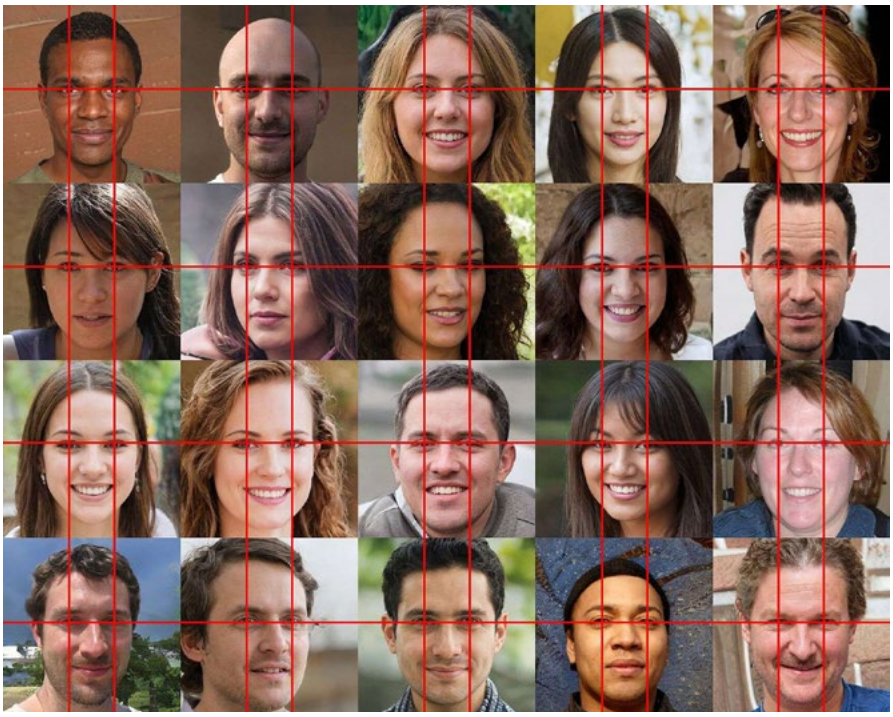
Thispersondoesnotexist.com: Rostros de personas inexistentes con una impresionante semejanza a humanos reales

Una plataforma web que facilita la creación de rostros falsos con IA es thispersondoesnotexist.com, a través de la cual se crean imágenes realistas de individuos ficticios. Este tipo de imágenes y otras de gatos imaginarios e incluso caricaturas que también pueden crearse con IA, pueden ser potencialmente usadas en operaciones de influencia en línea con las que se propaga desinformación, propaganda o contenido de odio.

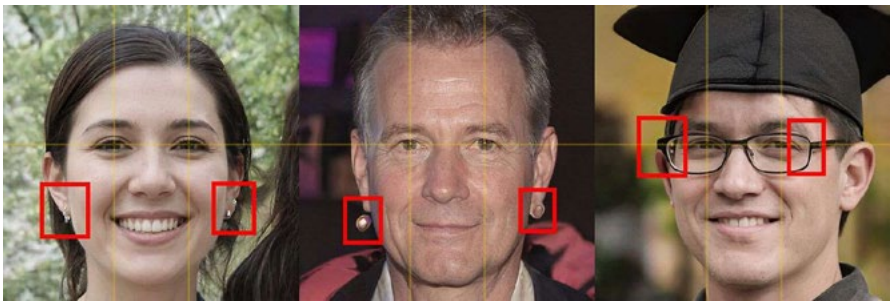
Cómo identificar rostros generados con IA

Debido al rápido avance de la tecnología generativa, detectar rostros creados con IA es cada vez más difícil. Sin embargo, en ocasiones es posible gracias a ciertas características distintivas, especialmente cuando se observan patrones coordinados en grupos de cuentas. Aquí hay algunas técnicas que pueden servir para identificar rostros generados sintéticamente:

hacer búsquedas como un experto

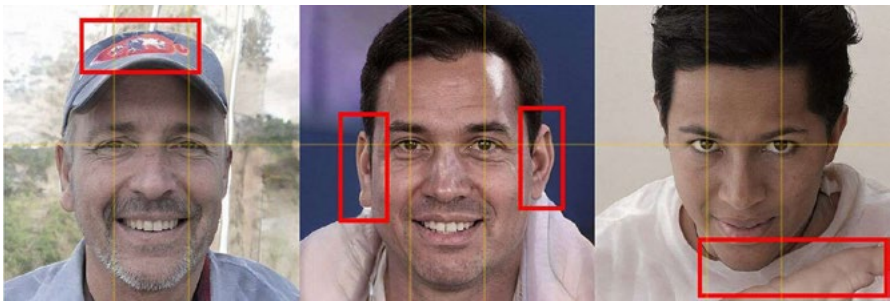


1 **Coincidencia en la posición del iris:** Los rostros generados por IA a menudo muestran una coincidencia en la posición del iris en múltiples imágenes. Este patrón es una pista común y útil al rastrear lotes de cuentas con fotos (posiblemente) creadas con IA.

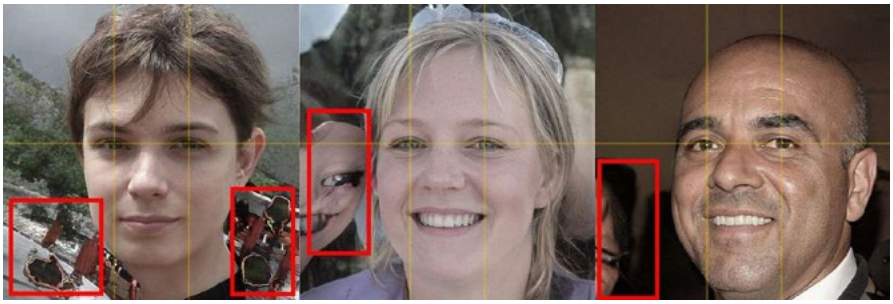


2 **Detalles asimétricos:** Los objetos como zarcillos, aretes, piercings y anteojos pueden no reproducirse con precisión en retratos sintéticos. Las asimetrías en estos detalles pueden ser señales claras de imágenes generadas por IA.

hacer búsquedas como un experto



3 **Deformaciones y desproporciones:** Se pueden identificar rostros con partes poco definidas o desproporcionadas, como partes de la cara sin forma definida o ropa con letras y logos ambiguos, a pesar de tener alta resolución.



4 **Distorsiones en partes distintas al rostro:** A veces, segmentos como caras, brazos u ojos pueden aparecer deformados. Además, se pueden encontrar glitches o errores digitales en áreas no relacionadas al rostro.

5 **Fondos distorsionados o inconsistentes:** Los rostros generados por IA pueden presentar fondos incoherentes, con manchas o patrones abstractos que no se corresponden con fondos reales.

hacer búsquedas como un experto

6 **Perfección fotográfica aparente:** Aunque algunas imágenes pueden tener errores y glitches, muchas creadas con IA tienen alta calidad, encuadre e iluminación. Esto, junto con fondos de estudio neutros, puede indicar su origen artificial. Además, cuando se solicita una imagen «hiperrealista», es común que las GANs añadan sombras pronunciadas y texturas aterciopeladas para lograr una apariencia fotográfica detallada. Sin embargo, todas estas características hiperrealistas pueden ser un indicio de que la imagen ha sido generada artificialmente.

hacer búsquedas como un experto

Inteligencia Artificial y videos

El crecimiento veloz de la Inteligencia Artificial ha iniciado una nueva era en la producción de contenido visual, donde esta tecnología es empleada tanto por profesionales como por actores desinformantes y maliciosos en todo el mundo. Esta tendencia global no ha excluido a Venezuela, donde la sociedad civil ya ha visto ejemplos claros del potencial para engañar en redes sociales de esta nueva tecnología. Sin embargo, conocer estas tácticas desinformativas y familiarizarse con las características que definen estas producciones manipuladas resultan fundamentales para contrarrestar su uso para distorsionar la información.

Falsos presentadores de noticias generados con IA

Una táctica que ha despertado preocupación en toda la región, es la proliferación de contenido *deepfake* con el que se simula la apariencia y la voz de personas reales.

Un caso elocuente es «House of News Español», un **falso canal de noticias en YouTube que usaba presentadores generados con IA**. El video en particular, sobre la recuperación económica de Venezuela en el Carnaval 2023, se destacó por su amplia circulación, que fue potenciada con el pago de publicidad en YouTube por parte de los creadores del canal.

Sin embargo, esos presentadores nunca existieron en la realidad; eran avatares digitales creados utilizando la plataforma synthesia.io. La herramienta permite crear avatares que hablen en múltiples idiomas y reciten textos de forma fluida, según lo instruido por cualquier usuario suscrito al servicio.

hacer búsquedas como un experto

Esta técnica no solo facilita la producción de contenido engañoso, sino que también representa un obstáculo considerable para la identificación de desinformación y propaganda encubierta.

Cómo detectar a un presentador generado por IA

Los siguientes son algunos indicadores que pueden ayudar a identificar la presencia de un presentador generado por IA:

1. **Patrones en los movimientos faciales que mantienen un eje fijo:** Uno de los signos reveladores de un presentador virtual es la rigidez en los movimientos faciales. A menudo, los avatares generados por IA pueden mostrar una falta de naturalidad en la variabilidad y expresividad de los gestos faciales, lo que resulta en movimientos que parecen estar restringidos a un patrón preestablecido.

2. **Fallas en la sincronización de la voz con los movimientos de los labios:** La sincronización entre la voz y los movimientos de los labios es un desafío común en la generación de presentadores virtuales. Si se observan discrepancias notables entre lo que se dice y cómo se mueven los labios, es posible que estemos frente a un video con un presentador generado por IA.

3. **Movimiento poco natural del cabello o de los accesorios como solapas de la camisa, cabello o zarcillos:** Los detalles físicos, como el cabello y los accesorios, suelen ser difíciles de simular de manera realista en los presentadores generados por IA. Si se detectan movimientos de cabello o de accesorios que parecen poco naturales o que

hacer búsquedas como un experto

carecen de realismo en su dinámica, esto podría indicar la presencia de una creación artificial.

4 **Hablar poco fluido, entonación monótona o ausencia de acentos locales:** La expresión vocal también puede proporcionar pistas sobre la autenticidad de un presentador virtual. Los avatares generados por IA pueden tender a mostrar una entonación monótona o una falta de fluidez en su discurso, ya que la replicación de la variabilidad y riqueza del lenguaje humano es un desafío complejo para las tecnologías actuales. Comparar el acento de la persona que habla con el usado por la población local, es también una manera de evaluar si se trata de un avatar.

TERCERA PARTE

Búsquedas avanzadas

03

03

TERCERA PARTE

Búsquedas avanzadas

Geolocalización: Investigando lugares

¿Cómo saber en qué lugar fue tomada una fotografía o un video? Una de las técnicas más útiles para determinarlo es la **geolocalización**.

La Real Academia Española define geolocalizar como la forma de «Determinar la ubicación geográfica de alguien o de algo valiéndose de medios técnicos avanzados, como el GPS». Es uno de los procedimientos más importantes para determinar el contexto original en el que se tomó una foto o un video viral y determinar si ha sido descontextualizado. Sin embargo, **muchos creen —erróneamente— que geolocalizar es un proceso complejo** que requiere de recursos tecnológicos que no están al alcance de todos, cuando, en realidad, existen técnicas sencillas y plataformas gratuitas que ayudan con el proceso.

En internet existen varios servicios en los que se puede acceder a mapas satelitales y, en algunos casos, observar de forma simplificada calles, centros urbanos, carreteras y otras imágenes que pueden ayudar a determinar si una foto o un video fue tomado en determinado sitio, sin tener que ir directamente al lugar.

Esta es una lista de los principales servicios de geolocalización disponibles que se pueden usar para investigar contenido multimedia:

búsquedas avanzadas

maps.google.com
earth.google.com

Google:

El rey de los buscadores también es un aliado a la hora de geolocalizar objetos o personas. A través de sus aplicaciones Maps y Earth se ofrecen imágenes de mapas desplazables, así como fotografías por satélite del mundo. Recomendamos descargar la versión de escritorio de Google Earth para mejores búsquedas.

yandex.com/maps

Yandex:

Este buscador ruso cuenta con una aplicación de mapas que ofrece imágenes desplazables, así como fotografías por satélite del mundo e incluso la ruta entre diferentes ubicaciones o imágenes a pie de calle.

bing.com/maps

Bing:

Microsoft Bing Maps es una web de mapas creada por Microsoft para su buscador Bing. Su principal competidor es Google Maps. En Windows 10 viene preinstalada con el nombre de Windows Maps.

apps.sentinel-hubs.com/sentinel-playground

Sentinel Hub:

Es un motor para procesar petabytes de datos satelitales. Esta plataforma SIG se basa en la nube para la distribución, gestión y análisis de imágenes satelitales. En ella se pueden combinar imágenes tomadas con diferentes bandas espectrales, que pueden ayudar a investigar incendios o derrames petroleros.

búsquedas avanzadas

Mapillary

mapillary.com

Mapillary:

Es un servicio de vistas de caminos por crowdsourcing o colaboración abierta. El servicio ofrece varias modalidades de captura de fotos, vía caminata, movilidad (a bicicleta o a motor) y fija.

¿Cómo se investiga en redes sociales?

Para investigar en redes sociales es importante considerar las opciones de búsqueda interna que ofrece cada servicio y, en algunos casos, el uso de los operadores, que permiten realizar búsquedas y hallazgos de información de forma más precisa.

Búsquedas en X/Twitter: con operadores

X (anteriormente conocida como Twitter) es una red social importante para la búsqueda de información, principalmente noticiosa, por lo que el uso de los operadores para abordar temas que son o hayan sido noticia a través de esta red puede arrojar muy buenos resultados.

A continuación presentamos los principales operadores de búsqueda que se usan en X/Twitter:

ESCANEA EL QR O HAZ CLICK PARA ACCEDER A LA APP CAZADORES

búsquedas avanzadas

Operador de Búsqueda	Descripción
"fútbol en vivo"	Contiene la frase exacta "fútbol en vivo".
amor OR odio	Contiene "amor" u "odio" (o ambos).
cerveza -raíz	Contiene "cerveza" pero no "raíz".
#vacaciones	Contiene el hashtag "vacaciones".
from:elpais	Muestra los posts enviados desde la cuenta de X/Twitter "elpais".
list:NASA/astronauts-in-space	Enviado desde una cuenta de X/Twitter en la lista de la NASA de astronautas en el espacio.
from:interior	Muestra posts de la cuenta de X/Twitter «interior».
to:NASA	Dirigido a la cuenta de X/Twitter de NASA.
@NASA	Menciona a la cuenta de X/Twitter de NASA.
gracioso filter:links	Contiene el término "gracioso" y enlaces a un URL.
mascota url:amazon	Contiene URLs que redirigen a Amazon y el término "mascota".
superheroe since:2015-12-21	Contiene el término "superhereo" y fue enviado desde el 21 de diciembre de 2015 en adelante.
mascota until:2015-12-21	Contiene el término "mascota" y fue enviado hasta el 21 de diciembre de 2015.

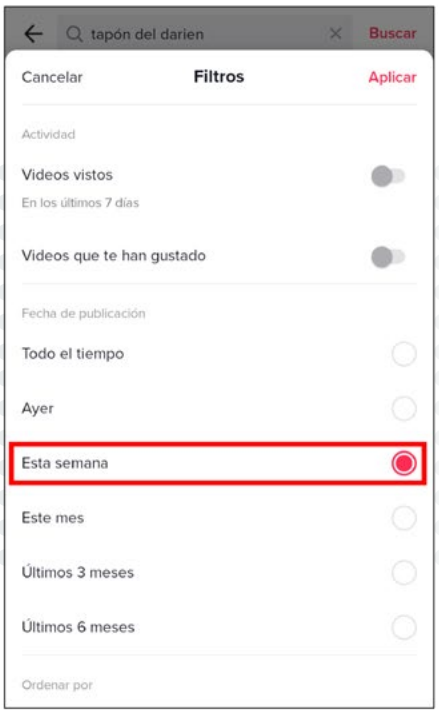
EN LA PESTAÑA «OPERADORES DE BÚSQUEDA» DE LA APP CAZADORES MANTENEMOS UNA LISTA ACTUALIZADA DE LOS OPERADORES DE BÚSQUEDA DE LAS PRINCIPALES PLATAFORMAS OSINT

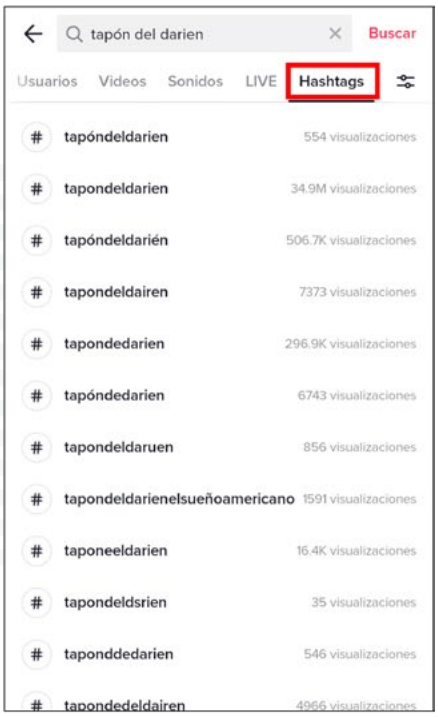
ESCANEA EL QR O HAZ CLICK

búsquedas avanzadas

Búsquedas en TikTok con etiquetas y filtros

TikTok cuenta con una interfaz muy limitada, que obliga a usar distintas pestañas para buscar vídeos y sonidos. También se pueden emplear filtros para organizar los resultados de búsqueda por su tiempo de publicación y etiquetas de distintos temas.





Búsquedas en Instagram con etiquetas y herramientas externas

Instagram es una red social muy limitada para la búsqueda de información, sin embargo existen varias herramientas que se pueden emplear dependiendo de lo que desees conseguir.

En la página web de Cazadores puedes encontrar la publicación ¿Cómo buscar datos en Instagram? en la que se muestran varias herramientas para búsquedas de publicaciones conociendo su lugar

búsquedas avanzadas

y fecha, búsquedas de hashtags más adecuados, de usuarios y de imágenes con mayor resolución, también hay técnicas para descargar historias antes de que desaparezcan y otras que sirven para obtener la mayor información posible en esta plataforma. Sin embargo, **CrowdTangle** –desarrollada por Meta– es la mejor opción para realizar investigaciones OSINT en Instagram, pero su acceso es limitado a periodistas y aliados de Meta.

Búsquedas en Facebook con WhoPostedWhat.com

En el caso de Facebook también puede usarse CrowdTangle, aunque hay una herramienta externa a la plataforma que es gratuita y accesible, llamada WhoPostedWhat **que suele entregar muy buenos resultados de búsqueda**. Esta herramienta permite buscar publicaciones con palabras claves claves en días, meses o años específicos.

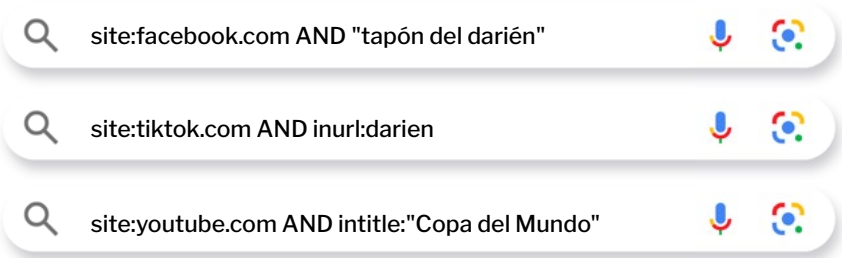
búsquedas avanzadas

Búsquedas en YouTube con filtros

La búsqueda de información en YouTube también está restringida al uso de los filtros, que ayudan a organizar los resultados de mayor a menor relevancia, por fechas de publicación y tiempo de duración del video, entre otras opciones.

Búsquedas en redes sociales desde Google usando operadores

A través de Google también se pueden hacer búsquedas usando operadores avanzados en algunas redes sociales como Facebook, TikTok, YouTube y X/Twitter, usando el operador «`site:`», combinándolos con otros términos y operadores. Abajo, un ejemplo de búsquedas en redes sociales desde Google, usando operadores:



CUARTA PARTE

Cuando la información tiene intención de daño

04

04

CUARTA PARTE

Cuando la información tiene intención de daño

La desinformación consiste en el uso de **información falsa o engañosa, generada para dañar deliberadamente a un objetivo**, pero no está limitada al uso de bulos. En operaciones de influencia, se hace uso de una variedad mucho más amplia de contenidos (incluyendo propaganda y datos verídicos) y tácticas para tergiversar la información en línea y alinear narrativas con una campaña comunicativa.

A menudo, los engaños más persistentes provienen de la repetición constante de narrativas, simplificaciones de la realidad que oscilan entre desinformación y propaganda.

Por lo anterior se hace crucial examinar las falacias lógicas prevalentes en estas operaciones, además de las tácticas principales empleadas para manipular conversaciones en redes sociales.

P. E. N. S. A. D. O. para manipular: el monopolio de la propaganda

El aparato de propaganda se nutre de recursos de argumentación en un **esquema pensado para reforzar sus narrativas** y minar los principios democráticos. En tiempos de storytelling, conviene conocer esas fórmulas.

En Cazadores identificamos como «cartas de monopolio» a siete técnicas y falacias lógicas que son usadas recurrentemente por propagandistas y actores desinformantes:

cuando la información tiene intención de daño

Carta 1: Provocación

Relacionada con la falacia «*cui bono*», consiste en culpar a adversarios de provocar incidentes en los que no están involucrados, porque supuestamente «los beneficia». Acusar a alguien de haber «provocado» algún hecho, pretende crear un elemento de incertidumbre y tentar al público a buscar «versiones alternativas» de los hechos. Es otro truco retórico barato para ocupar el espacio informativo», explica EUVsDisinfo, en su serie *Modus Trollerandi.*

En todos los casos, el objetivo de la provocación es sembrar dudas y disipar la atención de la audiencia.

Leer más aquí

Carta 2: Espantapajaros

El espantapájaros es un dispositivo retórico donde se atacan opiniones o ideas que nunca han sido expresadas por el oponente y que no están relacionadas con un argumento principal, que nunca es refutado directamente por el interlocutor. Está relacionada con la falacia del hombre de paja o del espantapájaros *(straw-man)*.

Leer más aquí

Carta 3: Negación

Entre los métodos desviadores, la negación es posiblemente el favorito. Con este método, el propagandista descarta cualquier evidencia o argumento documentado, negándolo parcial o totalmente. A veces, se hace uso de la falacia *ad*

cuando la información tiene intención de daño

ignorantiam, afirmando –erróneamente– que la ausencia de una prueba es prueba de su ausencia.

Leer más aquí

Carta 4: Sarcasmo

Según la Real Academia Española, la palabra sarcasmo significa «burla sangrienta, ironía mordaz y cruel con que se ofende o maltrata a alguien o algo». Como recurso de argumentación, su resultado tiende a ser humorístico, burlesco, o bien como una forma de crítica o censura respecto de algo. En ocasiones se usa el término «*Hahaganda*» para identificar a piezas de propaganda creadas con tono satírico.

Leer más aquí

Carta 5: Ataque

A este recurso de argumentación también se le conoce como falacia *ad hominem* o de Ataque Personal, ocurre cuando se ataca a la persona que emite un argumento, resaltando una supuesta incapacidad moral o o contradicciones de la parte acusatoria. De esta forma se desacredita lo que ha dicho por su persona para evitar que sus argumentos sean considerados.

Leer más aquí

cuando la información tiene intención de daño

Carta 6: Desviación

La táctica del *Whataboutismo* o del «*y tu también*», consiste en responder a una acusación o pregunta difícil formulando una contraacusación dirigida al emisor. En lugar de abordar o refutar el argumento original, esta técnica desvía la atención señalando que alguna persona o entidad vinculada al emisor también ha cometido actos similares en el pasado, con el objetivo de desacreditarlo.

Leer más aquí

Carta 7: Ofuscación

Esta última carta de la propaganda se basa en la premisa del ministro de propaganda nazi, Joseph Goebbels, quien dijo que «*si repites una mentira con la suficiente frecuencia, la gente e incluso tú la creerán*». En ocasiones, campañas de *phishing* o de estigmatización a periodistas, se basan en la repetición reiterada de un argumento o de un ataque por largos periodos de tiempo, lo que afianza una matriz de opinión errónea en el público objetivo: «si el río suena, es porque piedras trae». Está relacionado con la falacia lógica *ad nauseam*.

Leer más aquí

cuando la información tiene intención de daño

Manipulación informativa digital o «galería de trucos sucios»

En internet se hace uso de un sinnúmero de tácticas para manipular la conversación en línea, en ocasiones en el contexto de **operaciones de influencia** con las que se busca desinformar o difundir propaganda de forma sistemática. Además de estudiar los mensajes que se difunden, es importante evaluar la autenticidad y credibilidad de los canales por los que son compartidos, que en ocasiones son **activos digitales falsos** con los que se intenta simular, de forma encubierta, a usuarios o medios de comunicación.

Cuentas falsas: Son un tipo de activos digitales falsos creados y usados para que parezcan ser operados por usuarios o por grupos legítimos, aunque realmente son operados como marionetas para publicar desinformación, propaganda o contenido de odio de forma encubierta. Este tipo de cuentas anómalas se pueden dividir en varias categorías, entre ellas:

1. **Cuentas bot:** Funcionan mediante la automatización de ciertas tareas, generalmente la publicación de contenido, utilizando programas o *scripts* especializados. A menudo se asocian con la generación de contenido no deseado o *spam*, y suelen exhibir patrones de alta actividad centrados en temas específicos. Aunque no es una regla, estas cuentas pueden mostrar características como el anonimato, la ausencia de una imagen de perfil legítima, nombres de usuario que generan desconfianza, y una tendencia a compartir o validar contenido de otras cuentas con objetivos similares.
2. **Cuentas similares a bot:** Son cuentas que, si bien poseen una o varias características de cuentas

cuando la información tiene intención de daño

bot, no necesariamente están automatizadas, sino que son operadas por personas. En ciertas operaciones de influencia se prefiere la generación de *spam* con cuentas no automatizadas porque aparentan ser «menos falsas», evitando ser detectadas fácilmente por las redes sociales.

3. **Trolls:** Cuentas anómalas que generan contenidos polémicos y controversiales, para provocar reacción emocional o afianzar matrices de opinión. Según el Instituto Letón de Asuntos Internacionales, pueden ser de dos tipos: ***trolls tradicionales***, operados por usuarios espontáneos que publican comentarios polarizantes e intentan generar discusiones en línea, y ***trolls políticos***, creados en el contexto de operaciones de influencia, específicamente para difundir desinformación, propaganda o contenido de odio en redes sociales y defender o atacar a algún actor o partido político, marca o campaña.

Falsos noticieros: En ocasiones se crean cuentas de redes sociales o portales web que comparten contenido aparentemente noticioso, pero que no son medios reales, sino activos falsos que difunden desinformación y propaganda encubierta. Parte de su contenido puede ser copiado de medios de comunicación reales, que se intercala con contenido de menor calidad como memes, videos virales y contenido desinformativo o propaganda.

En algunos casos, falsos noticieros ofrecen servicios publicitarios encubiertos, que permiten publicar cualquier contenido –incluso desinformación– por un precio. También se ha detectado la creación de lotes de falsos medios tanto en redes sociales

cuando la información tiene intención de daño

como páginas web, en el contexto de operaciones de influencia desplegadas en épocas electorales, de alta conflictividad social o para promocionar positivamente a una persona o una causa.

Phishing: Es un conjunto de técnicas de ingeniería social con las que se busca engañar a víctimas al hacerse pasar por personas, empresas o servicios de confianza, para que hagan click en enlaces o abran documentos maliciosos, con el fin de manipularla o que realice las acciones deseadas por alguien. Puede tener distintos objetivos como obtener datos personales de los contactos de la víctima, hasta hacer ciberespionaje a periodistas.

Spam: Es un contenido repetitivo que se usa para manipular en redes sociales. Se le conoce como *contenido basura*, porque envía masivamente información que no fue solicitada y que puede tener fines publicitarios o hacer que contenido propagandístico o desinformativo parezca más popular de lo que realmente es.

Astroturfing: Es una técnica de marketing digital que consiste en activar de forma coordinada a una red de cuentas de redes sociales, falsas o no, para que impulsen una campaña publicitaria, un rumor o una matriz de opinión en redes sociales y hacerla ver como un tema del que se habla de forma espontánea.

Brigading: Es un término empleado por Meta para referirse a la acción de redes de cuentas —que pueden ser falsas o no— para acosar a personas u organizaciones mediante comentarios o respuestas dirigidos de forma organizada y coordinada. Se considera un tipo de *cyberbullying* que puede tener como fin generar miedo, aumentar el costo

ASTROTURFING ES TAMBIÉN UNA MARCA DE GRAMA ARTIFICIAL EN ESTADOS UNIDOS, Y PRECISAMENTE ESTE TÉRMINO SE EMPIEZA A USAR EN EL CAMPO DE LA DESINFORMACIÓN COMO UNA ALUSIÓN DEL EFECTO DE PERFECCIÓN DE LA GRAMA ARTIFICIAL QUE SE REPLICA EN REDES EN LAS CONVERSACIONES DE LAS REDES COORDINADAS, QUE PUEDEN GENERAR UN CONTENIDO QUE PARECE REAL PERO QUE REALMENTE ES INAUTÉNTICO.

cuando la información tiene intención de daño

de publicar contenido noticioso sobre un tema o perjudicar la reputación de periodistas o activistas.

Manipulación de encuestas en redes sociales: Una de las razones por las que las encuestas en medios de comunicación o redes sociales suelen ser poco confiables es su ausencia de valor porque están dirigidas a un público muy concreto. «No son muestras representativas de la población», explica Endika Núñez, analista de datos, en un artículo publicado en Maldita.es.

Cualquier encuesta publicada en redes sociales suele ser contestada por seguidores y usuarios cercanos a quien la generó, operando como una comunidad más o menos cerrada con puntos de vista u **opiniones que no representan el universo** de opiniones en toda la red social.

Algunas encuestas son herramientas de manipulación destinadas a persuadir a los lectores que pueden ser manipuladas fácilmente por redes de astroturfing e incluso redes de cuentas bot. Su influencia se amplifica cuando medios prestigiosos las difunden como noticias, sin verificar cómo fueron creadas o quién está detrás de ellas.

QUINTA PARTE

Operaciones de influencia

05

05

QUINTA PARTE

Operaciones de influencia

En el entramado de la desinformación es común escuchar hablar de las operaciones de influencia y de ciberincidentes. Aunque pueden parecer conceptos muy parecidos, no son sinónimos y su principal diferencia radica en sus objetivos finales.

Attribution.news describe a las **operaciones de influencia** como incidentes digitales basados en engaños que se despliegan en internet con el objetivo de persuadir a un público específico—en otras palabras, «*hackear mentes*»—para manipular su forma de pensar. Por su parte, los **ciberincidentes** también son incidentes digitales basados en engaño, pero que buscan «*hackear equipos*» (sistemas, celulares, servidores, páginas web, etc.), pudiendo comprometer la seguridad de sistemas y violar leyes o políticas, llegando incluso a interrumpir o destruir dichos sistemas.

En la práctica, cada operación de influencia y ciberincidente se estudia con metodologías similares, generalmente identificando las **tácticas**, **técnicas** y **procedimientos** usados. En ocasiones se despliegan operaciones de influencia y ciberincidentes de forma simultánea o la ejecución de uno puede allanar el camino para el despliegue del otro en una etapa posterior.

operaciones de influencia

Hablemos sobre operaciones de influencia

Con un objetivo final claro –*hackear mentes* a favor o en contra de una persona, grupo de personas o narrativas–, las operaciones de influencia son usadas para inundar el ambiente digital con ideas y emociones que condicionan las relaciones y procesos de pensamiento de otros usuarios para producir comportamientos afines con sus objetivos.

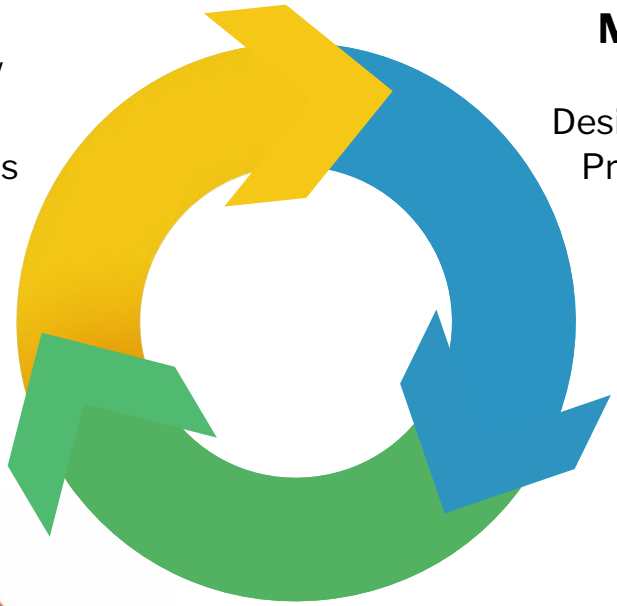
La configuración de una operación de influencia puede simplificarse en tres puntos indispensables: la estrategia, el mensaje y los activos.

Estrategia: Toda operación e influencia requiere de una planificación previa mínima, un diseño que permitirá alcanzar un objetivo específico. Antes de ser desplegada debe superar una fase inicial de planificación en la que se traza una «hoja de ruta» en la que se definen un mínimo de tácticas que serán usadas, como el impulso de etiquetas en redes sociales, el impulso coordinado de información mediante cuentas –falsas o no– o el empleo de *spam* para inundar con mensajes específicos los espacios digitales.

Mensaje: En una operación de influencia se diseñan mensajes que esperan llegar a la mente de la persona y cambiar su pensamiento sobre el tema que busca impulsar. El contenido compartido puede ser desinformativo, propaganda, rumores e, incluso, información verdadera, todo lo anterior con intención de generar daño a un adversario o favorecer la imagen de un aliado o cliente. El mensaje puede ser generado en forma de noticias, contenido audiovisual, memes, falsa publicidad, caricaturas y también con el uso de inteligencia artificial.

operaciones de influencia

Activos: es todo aquel vector digital o físico encargado de transmitir los mensajes. Pueden ser activistas, medios de comunicación, militantes de un partido, influenciadores, periodistas, funcionarios de seguridad (militares, policías) e incluso espontáneos cuyas ideas comulgan con lo que se quiere proyectar a través de la operación de influencia. También pueden usarse activos digitales falsos como noticieros falsos en redes sociales o en la web, cuentas bots, trolls, redes de astroturfing o falsos influenciadores reclutados para transmitir los contenidos.



operaciones de influencia

En medio de una operación de influencia también se pueden usar distintas estrategias para «*hackear mentes*»:

1. **Publicidad oscura:** promover «propaganda negra» o de «falsa bandera», es decir, pagar o promocionar contenido con ideas que aparentemente provienen de un bando político, pero que en realidad buscan dividir a ese mismo sector.
2. **Efecto rebaño:** promover grupos con ideas afines o contrarias que «pastoreen» al individuo y lo obliguen a unirse a un bando y descartar el otro, simplificando sus diferencias.
3. **Ataque en manada:** silenciar opiniones desfavorables atacando al emisor para que sienta que no tiene apoyo y que apoyarlo sea más difícil para otros usuarios.
4. **Promover las burbujas de opinión:** incentivar la uniformidad de pensamiento, para lograr que un público objetivo piense que investigaciones o contenido basadas en hechos no son más que mentiras.

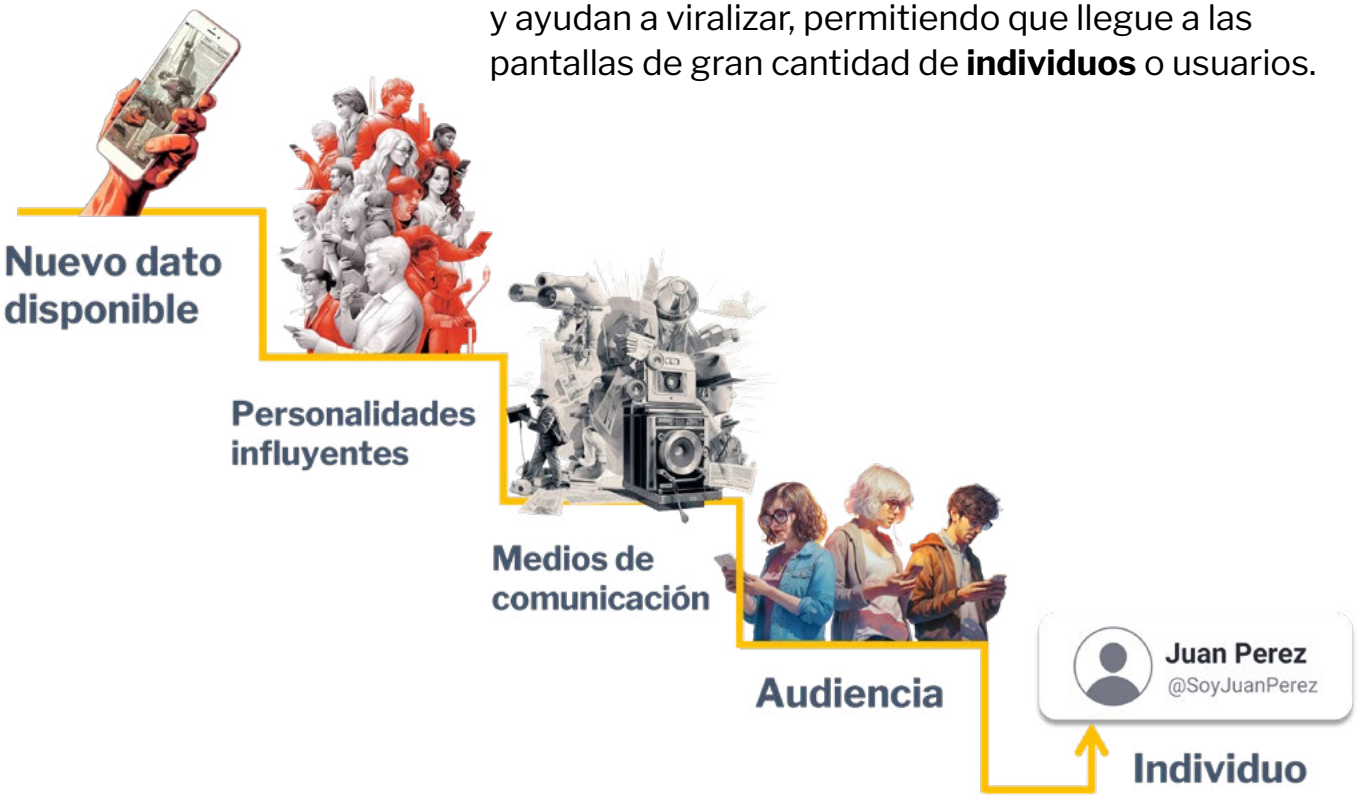
Lavar el origen de la desinformación

Así como en el mundo criminal se lava el dinero proveniente de negocios ilegales al pasarlos por un negocio legal, en una operación de influencia se busca «lavar» el origen de la desinformación cuando proviene de fuentes poco confiables, para lograr que **un actor ajeno a la operación y con validez en la comunidad transmita el mensaje** y lo valide como cierto.

Para entender este proceso se debe conocer el proceso por el que regularmente viaja la

operaciones de influencia

información. Generalmente inicia con el origen de un **nuevo dato** o información, que es presentado por **personalidades influyentes** (un científico, un denunciante, un político, un líder de comunidad). Posteriormente, la información capta la atención de **medios de comunicación** que lo presentan a un público mayor y la noticia es consumida por una **audiencia** amplia en redes sociales, que la comparten y ayudan a viralizar, permitiendo que llegue a las pantallas de gran cantidad de **individuos** o usuarios.



FLUJO DE TRANSMISIÓN DE INFORMACIÓN DESDE LA FUENTE HASTA EL CONSUMIDOR FINAL, SUSCEPTIBLE A MANIPULACIÓN POR ACTORES MALICIOSOS EN CUALQUIER ETAPA.

Las operaciones de influencia pueden introducir bulos, rumores o distorsiones informativas en cualquiera de los anteriores eslabones. Por ejemplo, en una operación de influencia se puede recurrir a la **creación de medios de comunicación** falsos para distorsionar una narrativa o desinformar sobre un

hecho. A nivel de audiencias se puede manipular la conversación en línea a través de redes de **redes de astroturfing**, o pueden atacar la reputación de **personalidades influyentes**, como periodistas, para intentar hacer mella en su credibilidad ante otros medios de comunicación y audiencias.

El peor de los escenarios en el lavado de la desinformación de una operación de influencia ocurre cuando actores reales (medios de comunicación reales, periodistas, personalidades) comparten un rumor, desinformación o propaganda encubierta **como si fuera información real** y lo amplifican sin saber realmente la intencionalidad del contenido que comparten.

Cuando esto ocurre, los actores desinformantes tienden a citar como fuente del contenido malicioso **al último actor influyente que lo mencionó**. De esta manera, omiten intencionadamente que la información originalmente provino de fuentes no confiables, como falsos noticieros, cuentas bots o trolls. Este enfoque sirve para «lavar» efectivamente el origen cuestionable de la información, aludiendo a la lógica de «*no lo afirmamos **nosotros**, sino que es algo que **ellos** mismos están diciendo*».

Contención de operaciones de influencia por parte de redes sociales

Los recursos usados en un conflicto bélico van mucho más allá de los aviones de guerra, tropas y misiles. En la actualidad, el espacio informativo también es considerado por muchos gobiernos del mundo como un escenario de guerra adicional. El concepto de

operaciones de influencia

Guerra Híbrida fue utilizado por primera vez a principios de los años 2000 y se refiere a la **implementación de estrategias de confrontación** que no necesariamente incluyen combates convencionales de tipo militar.

En este tipo de conflictos, además de emplearse acciones militares convencionales, también se hace uso de recursos no convencionales y de lucha no-violenta, con técnicas que incluyen ataques a objetivos cibernéticos, ciberespionaje y también el despliegue de operaciones de influencia, que buscan persuadir o desinformar a un público objetivo, incluso en redes sociales. Cuando estas **operaciones de influencia** son impulsadas por estados o fuerzas militares, suelen considerarse **operaciones de información** por redes sociales e investigadores.

operaciones de influencia

EL AVANCE Y ALCANCE DE LAS OPERACIONES DE INFLUENCIA HAN LLEVADO A EMPRESAS COMO X/TWITTER Y META A CREAR POLÍTICAS PARA DETECTAR ESTAS REDES E INTENTAR CONFRONTARLAS EN SUS PLATAFORMAS.

Desde 2016, diversas plataformas sociales como X/Twitter y Facebook han tomado medidas para mitigar las operaciones de información patrocinadas por estados. Esta decisión se adoptó en respuesta a controversias como el escándalo de Cambridge Analytica y el desmantelamiento de redes de cuentas trolls gestionadas por la Internet Research Agency, una entidad rusa vinculada con Yevgeny Prigozhin, líder del Grupo Wagner, que intentó **manipular el discurso en línea** durante las elecciones presidenciales de Estados Unidos en 2016.

Luego del descubrimiento de estos abusos en línea, las plataformas han desarrollado reglas y políticas que buscan contener el *spam*, la manipulación de plataforma, el comportamiento no auténtico coordinado, la difusión de contenido de odio y otras ciberamenazas, emitiendo reportes periódicos sobre las operaciones desmanteladas.

Meta (empresa matriz de Facebook, Instagram y WhatsApp), publica reportes periódicos sobre lo que denomina «amenazas adversarias», que clasifica en cuatro grupos distintos:

1. **Comportamiento no auténtico coordinado (CIB):** Son esfuerzos coordinados para manipular el debate público para alcanzar un objetivo estratégico, en los que se hace uso de activos digitales reales o falsos (cuentas de Facebook e Instagram y páginas y grupos de Facebook) para impulsar **operaciones de influencia encubiertas**.
2. **Brigading:** Se trata de esfuerzos coordinados para acosar a usuarios en redes sociales –entre ellos periodistas, medios y organizaciones no gubernamentales– en un intento de intimidarlos

operaciones de influencia

y silenciarlos, a menudo a través de mensajes directos repetitivos o comentarios masivos en sus publicaciones. Esta amenaza se relaciona con operaciones de ***ciberbullying*** digital.

3 **Espionaje Cibernético:** Se refiere a esfuerzos encubiertos para recopilar inteligencia y/o comprometer dispositivos y cuentas en línea de personas y cumplir objetivos estratégicos. Esta amenaza se relaciona con operaciones de **ciberespionaje** y su actividad puede ocasionar **ciberincidentes**.

4 **Redes de reportes masivos:** Son esfuerzos coordinados para abusar de los sistemas de reportes de contenido supuestamente violatorio de políticas de redes sociales, con el fin de eliminar cuentas o silenciar contenido legítimo y no violatorio de políticas, pero incómodo para determinado actor. Está relacionado con **campañas de censura** en línea en contra de periodistas, medios, activistas y organizaciones no gubernamentales.

Fases de una Operación de Influencia

El término «*Kill Chain*» se usa tanto en el ámbito militar como en el de inteligencia para describir los pasos o fases de un ataque. En el contexto de **operaciones de influencia**, existen metodologías plasmadas en forma de *Kill Chains* que ayudan a identificar las etapas de una operación impulsada por amenazas adversarias, al mismo tiempo que permite clasificar las tácticas, técnicas y procedimientos específicos que han utilizado.

operaciones de influencia

Desde 2023, Meta ha adoptado una metodología o Kill Chain de 10 fases, presentada por el investigador Ben Nimmo y el Carnegie Endowment for International Peace, para estudiar amenazas adversarias neutralizadas en sus plataformas.

La metodología sirve como una guía para los investigadores, quienes deben categorizar las diferentes tácticas observadas, clasificándolas dentro de cada fase. De este modo, pueden documentar de manera exhaustiva las características de las operaciones de influencia y otras ciberamenazas desplegadas en redes sociales, permitiendo estandarizar su estudio e intercambiar información de forma más organizada entre miembros de las comunidades de investigación OSINT, inteligencia y los investigadores internos de redes sociales.

Las 10 fases definidas en el *Kill Chain* usado en los reportes de amenazas adversarias de Meta, son:

1. **Adquirir activos:** En esta fase se adquieren activos digitales o físicos que formarán parte de la operación. Incluyen actividades como creación o compra de cuentas falsas en redes sociales, falsos noticieros, e incluso tácticas más orgánicas como el reclutamiento de influenciadores o periodistas y la creación de alianzas con grupos ciudadanos.

2. **Enmascarar activos:** Se trata de disfrazar activos falsos, comprados o creados para brindarles mayor legitimidad y hacer creer que se trata de cuentas o medios reales que distribuyen información de forma espontánea y no coordinada.

operaciones de influencia

3 **Conseguir información:** Se rastrea información en fuentes abiertas o bases de datos cerradas disponibles, con el fin de planificar la estrategia de la campaña y microsegmentar la audiencia de forma efectiva.

4 **Coordinar y planificar:** Para probar comportamiento no auténtico coordinado, es imprescindible probar coordinación. Por ello, se hace uso de tácticas como coordinar a grupos de personas, campañas de astroturfing y redes de cuentas bot, coordinación en grupos de WhatsApp e impulso de etiquetas específicas en redes sociales.

5 **Probar defensas de la plataforma:** En medio de las operaciones de influencia se pueden realizar pruebas previas tanto de los activos a usar como del contenido que será impulsado, para disminuir la probabilidad que tienen de ser detectados automáticamente en redes sociales, que violen reglas o políticas o que sean reportados masivamente por usuarios como contenido tóxico. Así, evitan que pierdan su impacto en momentos claves como, por ejemplo, durante una campaña electoral.

6 **Evadir detección:** A través de técnicas como el cambiar las letras de algunas palabras, publicar *spam* no relacionado con la propaganda o desinformación que se impulsa («*spamuflaje*»), se intenta se intenta evitar que las redes sociales detecten las actividades maliciosas.

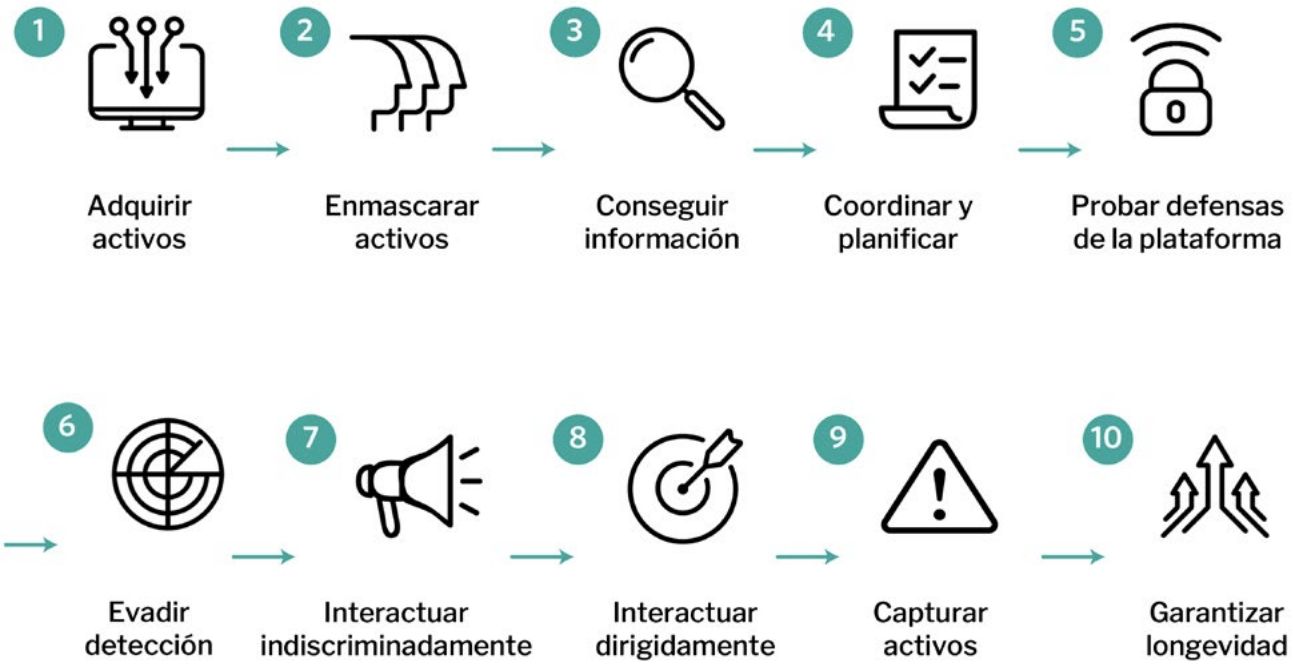
7 **Interactuar indiscriminadamente**: Compartir información o impulsar etiquetas de forma masiva, sin necesariamente dirigirlo a un público específico. De esta forma se intenta amplificar o hacer que un mensaje parezca más relevante o confiable de lo que realmente es.

operaciones de influencia

8 **Interactuar dirigidamente:** Compartir mensajes o etiquetar a periodistas, influenciadores o personas de interés que puedan ayudar a amplificarlo – inadvertidamente o no– y de esta forma lavar su origen y brindarle mayor credibilidad.

9 **Capturar activos:** Relacionado con el hackeo de activos digitales (cuentas, sitios web), que pueden formar parte de una operación de influencia, un ciberincidente o una campaña maliciosa.

10 **Garantizar longevidad:** Se relaciona con varias tácticas para lograr que una campaña perdure en el tiempo, incluso después de ser detectada, entre ellas cambiar nombres de empresas de comunicación política después de que su actividad ha sido expuesta, modificar nombres de cuentas troll, borrar tweets incriminatorios, etc.



FASES DE KILL CHAIN NIMMO/META PARA EL ESTUDIO DE AMENAZAS ADVERSARIAS

operaciones de influencia

Operaciones de influencia en contra de la prensa y de la sociedad civil venezolana

Los medios, periodistas y organizaciones de la sociedad civil en Venezuela son particularmente **susceptibles a ataques** en el ámbito de la información digital. A diferencia de actores políticos, que suelen tener equipos de apoyo, los profesionales de la información y activistas ciudadanos generalmente carecen de los recursos y el conocimiento para defenderse de tácticas de desinformación, manipulación en línea y ciberincidentes.

Desde el año 2019, Cazadores de *fake news* ha documentado varias decenas de **ataques digitales** en contra de medios, periodistas y organizaciones venezolanas. Clasificamos estos ataques en dos grandes grupos: ataques espontáneos y ataques coordinados.

Ataques espontáneos: Estos ataques suelen ser obra de usuarios individuales, periodistas, influenciadores o cuentas que actúan de forma individual, sin coordinarse con cuentas de otras redes existente o formar parte de operaciones de influencia específicas. La naturaleza aislada de estos ataques los hace particularmente **difíciles de contrarrestar**. Dado que no hay una amplificación coordinada de la información, las plataformas de redes sociales suelen interpretar estas acciones como comentarios amparados por la libertad de expresión –a menos que violen claramente políticas como divulgación de documentos personales o contenido de odio–, lo que evita, en muchos casos, que activen los protocolos para restringir o moderar dicho contenido.

operaciones de influencia

Ataques coordinados: Estos ataques provienen de redes de cuentas o noticieros que están vinculados en red. Un ejemplo claro son los equipos de **astroturfing**, contratados específicamente para llevar a cabo ataques dirigidos.

En Venezuela, se ha observado que líderes de grupos de astroturfing ofrecen abiertamente compensación económica en redes sociales para fomentar el uso de etiquetas coordinadas que difaman a periodistas, medios de comunicación, organizaciones y figuras políticas. También se ha observado que **falsos noticieros de Instagram** han publicado contenido coordinadamente para darle mayor credibilidad a

operaciones de influencia

un bulo o un rumor. En ocasiones, existen ataques coordinados que hacen uso de múltiples técnicas y se despliegan en varias plataformas simultáneamente. También tenemos el caso de las **cibertropas digitales** que pueden usar cuentas reales o falsas, alineadas con cuerpos policiales y militares, que buscan impulsar propaganda favorable y contrarrestar narrativas incómodas sobre su estructura y funcionamiento.

X/Twitter considera algunos ataques coordinados como violatorios de su Políticas en contra del *spam* y la manipulación de plataforma, debido a la generación de *spam*, cuentas falsas y coordinación para hacer que un tema parezca más importante de lo que realmente es. Meta, por su parte, puede identificar a estos ataques coordinados como campañas de brigading, reportes coordinados o Comportamiento no auténtico coordinado (CIB).

En ocasiones, es posible atribuir ciertas operaciones de influencia a estados o cuerpos de seguridad, considerándose **operaciones de información**. El caso más emblemático en Venezuela es la red «**Tuiteros de La Patria**», una red de *spam* en X/Twitter con la que se amplifican etiquetas propuestas a diario por el Ministerio del Poder Popular para la Comunicación e Información y que ha intentado ser controlada en varias ocasiones por la red social. Las cuentas que forman parte de esta red son manejadas por venezolanos comunes que, al menos hasta junio de 2023, recibieron pagos a través del Sistema Patria por participar en la operación de información, violando las políticas de X/Twitter en contra del *spam* y la manipulación de plataforma.

operaciones de influencia

En ocasiones, las operaciones de información van acompañadas por ciberincidentes que implican phishing, hackeos o difusión de virus de forma dirigida a periodistas o miembros de la sociedad.

Algunos casos de estudio

En Cazadores hemos logrado detectar varios ataques enmarcados en operaciones de información o de influencia en contra de periodistas que mantienen patrones comunes y usan tácticas relacionadas con desinformación.

Entre los más destacados podemos nombrar campañas de estigmatización, campañas de brigading, el uso de redes de noticieros falsos y los ataques trolls.

Campañas de estigmatización

Norbey Marín, presentador del programa de YouTube «Hasta que Caiga la Tiranía,» fue objeto de **dos campañas coordinadas de difamación** en Twitter el 8 y 9 de noviembre de 2021. Estas campañas fueron impulsadas por la misma red de astroturfing que previamente había atacado a figuras como como Juan Guaidó y el padre Arturo Sosa Abascal. Las etiquetas #NorbeyExtorsionador y #NorbeyProfugo circulaban junto a mensajes y memes difamatorios.

La difamación no se limitó a Twitter. El 9 de enero, los memes y mensajes difamatorios se publicaron en una reden una red de al menos cinco noticieros falsos en Instagram, conocida por Cazadores de *fake news* como **«La Fábrica de Desinformación.»** Esta red ha sido la fuente de varias decenas de desinformaciones y rumores difundidos desde 2019, dirigidas principalmente contra figuras políticas, periodistas y medios venezolanos.

operaciones de influencia

Tendencias coordinadas en Twitter

Falsos noticieros coordinados en Instagram

Para leer más sobre este caso puedes entrar aquí.

Campaña de Brigading

Sebastiana Barráez, una periodista venezolana de gran influencia especializada en temas militares, ha dedicado su carrera a publicar numerosos artículos sobre la presencia y las actividades de grupos guerrilleros y paramilitares colombianos en territorio venezolano. Su trabajo ha sido una fuente clave de información sobre estos complejos temas.

El 19 de septiembre de 2020, se produjo un incidente en el estado Apure que involucraba a la Fuerza Armada Nacional Bolivariana (FANB) y a grupos guerrilleros colombianos. Al día siguiente, la FANB emitió un comunicado en su cuenta oficial de prensa en el que se abstuvieron de referirse a los grupos involucrados como guerrilleros. Sebastiana Barráez criticó públicamente esta omisión a través de un tweet el 20 de septiembre, lo que desencadenó

operaciones de influencia

una serie de **ataques dirigidos** hacia ella en la red social durante los días 21 y 22 de septiembre, un ejemplo de ciberacoso o lo que algunos años más tarde, Meta comenzaría a llamar «**Brigading**».

Los tweets en contra de Barráez fueron publicados de forma coordinada por decenas de cuentas de funcionarios de distintos cuerpos de la FANB, además de una subred de cuentas falsas que fueron creadas pocos días antes del ataque, suspendidas luego de la denuncias realizadas por Cazadores de *fake news* en Twitter.

Para leer más sobre este caso puedes entrar aquí.

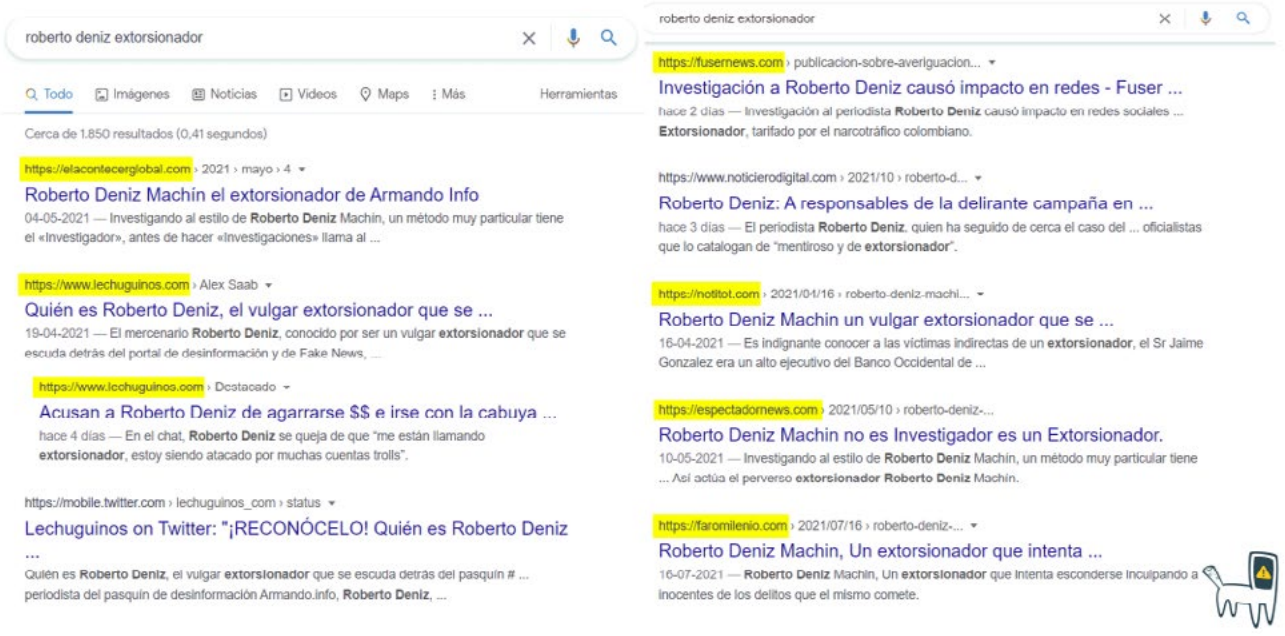
Redes de noticieros falsos en la web

Una empresa de marketing digital participó en la creación de una red de al menos **22 portales de noticias falsas** entre abril y junio de 2021. Estos sitios difamaron a los investigadores **Roberto Deniz** de Armando.info y **Alek Boyd** de Infodio.com, tildándolos de «extorsionadores».

operaciones de influencia

Esta red se inserta en una operación de influencia más amplia a favor del empresario Alex Saab, que ha estado en marcha en internet desde 2020. Al menos cuatro empresas de marketing digital en Venezuela y varios países africanos han estado implicadas en la campaña. Deniz y Boyd han reportado sobre una trama de corrupción que involucra a Saab, quien estuvo detenido en Cabo Verde hasta su extradición a Estados Unidos el 16 de octubre de 2021. La atribución de la campaña difamatoria se hizo mediante el uso de fuentes abiertas y públicamente accesibles.

El mismo día de la publicación del reporte sobre la red de falsos noticieros, la dirección de la empresa de marketing digital dio de baja a todos los falsos noticieros, argumentando que solamente había sido intermediaria en la creación de los sitios web y negando tener cualquier tipo de relación con la operación de influencia a favor de Alex Saab.



Para leer más sobre este caso puedes entrar aquí.

operaciones de influencia

LO QUE CREE EL LECTOR:

«VARIOS PORTALES RESEÑARON LA INFORMACIÓN, DEBE SER VERDAD»

Ataque troll contra El Pitazo y Provea

Durante marzo de 2021, un conjunto de **85 cuentas de Twitter** dirigió múltiples respuestas de rechazo hacia tweets publicados por el portal venezolano **El Pitazo** y la ONG **Provea**. Estas respuestas cuestionaban dos artículos que criticaban la actuación de las Fuerzas de Acciones Especiales (FAES), un comando de la Policía Nacional Bolivariana de Venezuela.

Usuarios venezolanos observaron que las cuentas involucradas en estos ataques eran sospechosas: en su mayoría creadas en febrero de 2021, con perfiles anónimos o con escasa información. Las cuentas parecían haber sido creadas específicamente para defender a las FAES.

La investigación reveló que estas cuentas formaban parte de una red de 85 cuentas falsas de Twitter creadas desde enero de 2021 con el objetivo de generar contenido en defensa de las FAES. Las cuentas intentaban simular apoyo popular para las FAES mientras criticaban a medios o entidades que cuestionan sus actividades. La actividad de la red troll violó la política de Twitter sobre *spam* y manipulación de la plataforma, siendo suspendida en su totalidad luego de que Cazadores de *fake news* publicara un reporte sobre la operación de influencia.

operaciones de influencia

Para leer más sobre este caso puedes entrar aquí

SEXTA PARTE

Rastreo de rumores en tiempos de crisis

06

06

SEXTA PARTE

Rastreo de rumores en tiempo de crisis

Periodistas, políticos, organizaciones de la sociedad civil y activistas a nivel global se encuentran frecuentemente en la mira de operaciones de influencia y campañas de estigmatización en línea. Actuar de manera rápida y precisa es crucial para **mitigar los ataques** y **minimizar su impacto.**

El Consejo de Derechos Humanos de la Organización de Naciones Unidas (ONU) advierte sobre los riesgos inherentes a la estigmatización. Según esta entidad, la estigmatización tiene como objetivo caracterizar a individuos o grupos como **«terroristas»**, **«enemigos del Estado»** u **«oponentes políticos»**, a menudo a través de autoridades estatales y medios de comunicación, para deslegitimar su trabajo y aumentar su vulnerabilidad a abusos y violaciones de derechos humanos.

La estigmatización en línea no solo manifiesta ataques a través de redes sociales, sino que también representa un serio **peligro para los derechos humanos de la persona afectada.** Ignorar estos ataques no es una opción viable dada la gravedad de las implicaciones.

Como un caso emblemático, la periodista Gauri Lankesh fue asesinada en Bangalore, India, el 5 de septiembre de 2017. Lankesh, quien había estado en el foco de campañas de odio digital, escribía sobre la propagación de la desinformación en India y había detallado cómo ciertos sitios web y actores políticos utilizaban noticias falsas

rastreo de rumores en tiempos de crisis

como armas. El ataque a Lankesh demuestra la extrema vulnerabilidad a la que están expuestas las personas objeto de estigmatización en línea.

Es importante para periodistas, medios y organizaciones que sospechan ser objetivo de campañas de estigmatización en línea:

1. **No ignorar el problema**: existen protocolos mínimos que se deben seguir ante ataques de este tipo, que deben ser activados de inmediato para documentar, recabar evidencias y, eventualmente, denunciar el ataque. Ignorar el problema no es una opción.
2. **Diseñar un plan de rastreo sistemático de rumores o ataques**: indistintamente de la gravedad o la seriedad de los ataques, es recomendable documentarlos ya que pueden dar pie al establecimiento de patrones de conducta que permitan atribuir los ataques recibidos.
3. **Reportar y pedir ayuda oportunamente**: organizaciones de la sociedad civil pueden ayudar a documentar, evaluar y denunciar ataques de estigmatización y ciberincidentes digitales. También podrían servir como puentes con redes sociales que permitan contener operaciones que violen reglas y políticas de las plataformas.

Existen diversas metodologías para llevar a cabo el **rastreo y el estudio de rumores**, que pueden resultar en desinformación, propaganda o incluso en contenido noticioso si se verifican. Uno de esos métodos es el detallado en los manuales «Managing Misinformation in a Humanitarian Context: Context & Case Studies», que conforman una metodología para rastrear rumores desarrollada por Internews.

rastreo de rumores en tiempos de crisis

Esta metodología puede aplicarse en diversos contextos: desde campañas de estigmatización contra líderes políticos y periodistas, olas de rumores durante situaciones de conflictividad política, social o humanitaria e incluso para realizar el seguimiento sistemático de ciberincidentes. Está compuesta por cinco etapas:

1 **Comprender el contexto y la comunidad**: Es crucial entender el ambiente en el que se desarrolla una operación de influencia, considerando las características del contexto. No es igual el tipo de ataques o rumores que pueden detectarse en un campamento de migrantes, que los ataques y ciberincidentes que se pueden registrar durante una campaña política o durante una ola de ciberataques dirigida a los sistemas de un hospital.

2 **Planificación del proyecto**: Se debe crear una tabla detallada de rastreo de incidentes, con categorías clave para el rastreo y la documentación de información, estableciendo prioridades, y buscando patrones de conducta. Esta tabla puede ser creada en Google Sheets o Microsoft Excel y contener columnas como el miembro del equipo encargado de registrar el incidente, su fecha, fuente y la red social donde se generó el rumor, el tipo de contenido y el nivel de urgencia para abordarlo.

3 **Recopilación de casos**: todos los incidentes monitoreados, indistintamente de que sean rumores, bulos o ciberincidentes, se registran en la tabla previamente diseñada para obtener una visión completa y detectar patrones durante el análisis global.

rastreo de rumores en tiempos de crisis

#	Encargado	Fecha	Fuente	Red social	Descripción	Tipo	Urgencia	Observaciones

EJEMPLO DE TABLA PARA REGISTRAR INCIDENTES O RUMORES DE ACUERDO A LA METODOLOGÍA PLANTEADA

4 **Análisis y diseño de respuestas**: El análisis de la información es una etapa crucial, y se deben priorizar respuestas a los casos más importantes, ya que abordar todo el contenido registrado en la tabla de rastreo podría ser inviable. Para determinar qué rumores o ataques requieren atención, es esencial hacerse las siguientes preguntas:

- ¿El rumor o ataque puede dañar a la comunidad o promover un comportamiento de riesgo?
- ¿Presenta un riesgo de seguridad para un grupo específico?
- ¿Puede poner en riesgo al personal o a proveedores/aliados?
- ¿Podría inhibir o prevenir el uso de servicios o el ejercicio de derechos?
- ¿Representa un riesgo reputacional?

rastreo de rumores en tiempos de crisis

Una vez se realiza el análisis y se tienen todos los datos para contrastar la información, es necesario diseñar un plan para la divulgación del contenido o para responder al ataque. Aquí hay algunas pautas importantes:

- **No todos los rumores o ataques requieren respuesta**. Se debe evaluar cada caso para determinar si responder es apropiado o si podría, de hecho, amplificar un rumor o ataque poco visible o importante.
- Las respuestas **no deben parecer defensivas o reactivas**.
- Los desmentidos o aclaraciones **deben ser más que simples negaciones**; deben, idealmente, explicar el origen del rumor o ataque y mostrar evidencia si es posible.

5 **Divulgación de contenidos y respuestas**: Una vez diseñado el plan de respuesta, es momento de ejecutarlo, divulgando las aclaraciones, desmentidos o denuncias pertinentes.

Epílogo

epílogo

Al concluir el último capítulo de esta recopilación de **conceptos, técnicas y metodologías para combatir la desinformación**, aspiramos a proporcionar al lector una visión más clara sobre una problemática que, si no es investigada ni controlada, puede afectar la salud de la democracia. **La desinformación, un actor silencioso pero potente**, se manifiesta en múltiples facetas en Venezuela y se replica en otras latitudes de América Latina, dejando cicatrices que llaman nuestra atención y lecciones que nos invitan a cambiar nuestra forma de consumir contenidos informativos.

Debido a los nuevos desafíos que plantea este fenómeno, surge la necesidad de fortalecer el pensamiento crítico en la sociedad civil. Como siempre solemos discutir con nuestros pares cazadores de noticias falsas –*ciudadanos, periodistas, activistas e investigadores*–, este es el momento para superar etapas, siendo el pensamiento crítico el único camino que tenemos para lograr contener la desinformación. **Los verificadores son inútiles si las personas no comienzan a comprender** lo que está ocurriendo. Y este no es un desafío que solo enfrentan los venezolanos, sino una realidad que reverbera en los corazones de nuestras naciones hermanas.

Vemos a este manual como una instantánea de nuestras experiencias estudiando el fenómeno de la desinformación y las operaciones de influencia en la región, basado también en lecciones aprendidas en otras latitudes y pocas veces explicadas en idioma español. Aspiramos a que sirva para **mejorar la resiliencia frente a estos problemas** en el seno de nuestra sociedad, partiendo de nuestra creencia que la desinformación, al ser desmenuzada, comprendida y explicada, deja de ser un fantasma para convertirse en un desafío tangible y superable.

epílogo

La dinámica de la desinformación es una batalla que perdurará por muchos años más, a pesar de los legítimos intentos de contención que puedan ser impulsados desde grandes plataformas de redes sociales o por cuerpos legislativos de nuestros países. No obstante, surge la esperanza de enfrentar de forma sostenible a este problema cuando reconocemos que **su punto débil son los esfuerzos investigativos** de periodistas y organizaciones que comprenden las dinámicas del fenómeno, dispuestas a desentrañar la verdad e invitar a que los ciudadanos se unan como un frente consolidado, bien informado sobre esas tácticas, con el fin de contrarrestar sus efectos.

La lucha contra la desinformación se halla, por tanto, en una encrucijada que nos exige elevar el costo de desinformar y engañar a nuestras sociedades. Un costo que debe resonar en los rincones de la conciencia colectiva y resaltar **la importancia de tener sociedades informadas y libres.**

Al cerrar estas páginas, el llamado es a **continuar trabajando conectados como un engranaje**, fomentando la búsqueda y presentación de evidencias, manteniendo encendida la llama del análisis crítico y exigiendo a distintos actores que actúen de la forma más transparente posible.

Porque el valor de la verdad, en una sociedad informada y libre, es incalculable.

Adrián González
Director Cazadores de Fake News

índice de conceptos, servicios y herramientas

índice de conceptos, servicios y herramientas

índice de conceptos, servicios y herramientas

Cazadores de **Fake News**
cazadoresdefakenews.info