Centro Universitario de la Defensa
en la Escuela Naval Militar

TRABAJO FIN DE MÁSTER

Anonimización, ocultación y eliminación de huella digital

Máster Universitario en Dirección TIC para la Defensa

ALUMNO: Ignacio Gómez Burgaz

DIRECTORES: Javier Vales Alonso

José González Coma

CURSO ACADÉMICO: 2023-2024

Universidade Vigo

Centro Universitario de la Defensa
en la Escuela Naval Militar

TRABAJO FIN DE MÁSTER

Anonimización, ocultación y eliminación de huella digital

Máster Universitario en Dirección TIC para la Defensa
Especialidad de Sistemas y Tecnologías de la Información /
Especialidad de Sistemas y Tecnologías de Telecomunicación

Universidade Vigo

RESUMEN

La necesidad y el uso de internet en el mundo actual nos sugiere analizar diversos métodos y acciones que tanto un usuario individual como una organización pueden llevar a cabo para navegar, obtener y gestionar información de manera anónima en un entorno digital. Este objetivo se plantea con el propósito de asegurar una comunicación y navegación anónimas y seguras eludiendo la detección, seguimiento y monitorización por terceros.

Desde la privacidad de cualquier usuario en el espacio virtual hasta la configuración de un entorno seguro en el contexto de una organización se exploran las medidas de seguridad para salvaguardar la privacidad y realizar actividades como la recopilación de información. Adicionalmente, se abordan las normativas jurídicas y consecuencias legales asociadas a la ley de protección de datos, la privacidad y el derecho al olvido.

Se proponen estrategias aplicables para el uso y la navegación no solo en ordenadores personales, sino también en dispositivos móviles e incluso dentro de la red de trabajo de una organización.

A lo largo del desarrollo del trabajo, se realizará un análisis detallado de las técnicas de anonimización y ocultación existentes, evaluando sus características y aplicaciones específicas. Este análisis se extenderá al estudio de las características de los diferentes navegadores y buscadores, explorando sus funcionalidades y las posibles huellas digitales generadas por los usuarios.

Finalmente, se abordará en detalle la huella digital generada por la actividad de navegación, incluyendo el registro en diversas plataformas, se examinará el funcionamiento de herramientas de seguimiento como cookies y supercookies, y se propondrán diversos métodos para evitar la monitorización por parte de herramientas y servicios de análisis que puedan estar integrados o incrustados en dispositivos y plataformas.

PALABRAS CLAVE

Anonimicidad, ocultación, huella digital, seguridad informática, privacidad digital

AGRADECIMIENTOS

La realización y terminación de este trabajo ha sido posible gracias a múltiples factores, razón por la cual me gustaría dedicar unas pocas palabras a toda esa gente que me ha apoyado, impulsado y animado a lo largo no solo de la elaboración de este trabajo sino en la consecución y terminación de este Master pese a todas las vicisitudes sufridas.

A mis compañeros de Master y profesores del CUD con los que hemos conseguido crear unos lazos de compañerismo y amistad que perdurarán por mucho tiempo además de las aventuras vividas.

A mis tutores del trabajo, por orientarme y darme las indicaciones correctas en los momentos en que las circunstancias me sobrepasaban.

Y por último y no por ello menos significativo, querría agradecer a mi familia todo por el trabajo extra que han realizado al tener que soportarme, aguantarme y a su vez y más importante, apoyarme durante las largas horas de trabajo

Espero que éste sea un documento interesante y de su agrado.

CONTENIDO

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

1 INTRODUCCIÓN Y OBJETIVOS

1.1 Contexto y relevancia del tema.

La evolución histórica de Internet se remonta al desarrollo de las redes de comunicación en el contexto de la Guerra Fría, que dio origen a la primera red de ordenadores llamada ARPANET y que luego se transformó en Internet al adoptar el protocolo TCP/IP. A finales del siglo XX, Internet se expandió por todo el mundo y se popularizó el uso de la World Wide Web (WWW), que permitió el acceso a una gran cantidad de información y servicios online. Internet ha supuesto un factor transformador tanto a nivel individual como social, pero también ha planteado nuevos desafíos y riesgos para la privacidad de los usuarios.

A raíz de este factor transformador, ha surgido, en línea con los desafíos anteriores y ahondando en el tema de la privacidad, la recopilación de datos. La recopilación de datos es el proceso de reunir y medir información de diversas fuentes con fines de investigación y análisis. En Internet, existen diferentes técnicas y herramientas para recopilar datos digitales, tanto de forma directa como indirecta, que pueden proporcionar un perfil detallado de los usuarios, sus preferencias, comportamientos y hábitos de consumo. La recopilación de datos puede tener beneficios como el respeto a la intimidad, el cumplimiento de la normativa, la mejora de la calidad de los datos o la innovación en el análisis de datos.

La anonimización, ocultación y eliminación de huella digital son conceptos relacionados con la privacidad y la seguridad en Internet. La huella digital es el conjunto de datos e información que dejamos al navegar por la red, como los sitios web que visitamos, los correos electrónicos que enviamos, las redes sociales que usamos y la información personal que compartimos. Estos datos pueden ser recopilados y utilizados por terceros con fines comerciales, publicitarios, de vigilancia o de ciberdelincuencia. Por eso, muchas personas quieren reducir o eliminar su huella digital para proteger su identidad, su reputación y sus derechos.

La anonimización es el proceso de eliminar o modificar los datos personales que permiten identificar a una persona, como el nombre, el correo electrónico, el número de teléfono o la dirección IP. La anonimización puede hacerse mediante técnicas como el cifrado, el enmascaramiento, la agregación o la generación de datos sintéticos. El objetivo es que los datos anonimizados no puedan ser asociados a una persona concreta, ni siquiera con técnicas de reidentificación.

La ocultación es el proceso de evitar que los datos personales sean recopilados o rastreados por terceros, como los proveedores de servicios de Internet, los motores de búsqueda, las redes sociales o las empresas de publicidad. La ocultación puede hacerse mediante técnicas como el uso de navegadores privados, el bloqueo de cookies, el uso de VPN o el uso de servicios que no requieren registro o que permiten el uso de seudónimos.

La eliminación de la huella digital es el proceso de borrar o solicitar el borrado de los datos personales que ya han sido recopilados o publicados en Internet, como los perfiles de redes sociales, las suscripciones, los servicios web o las cuentas de compra. La eliminación puede hacerse mediante técnicas como el uso de herramientas que rastrean y borran los datos personales en diferentes sitios web, el ejercicio del derecho al olvido o el uso de servicios que facilitan la baja o el cierre de cuentas.

Estos conceptos son importantes porque afectan a la privacidad y la seguridad de las personas en Internet, así como a sus derechos fundamentales. La anonimización, ocultación y eliminación de huella digital pueden ayudar a prevenir el robo de identidad, el acoso, la discriminación, el fraude, el espionaje o la manipulación. Sin embargo, también pueden tener limitaciones, riesgos o consecuencias negativas, como la pérdida de información, la dificultad de acceso a ciertos servicios, la responsabilidad legal o la falta de transparencia. Por eso, es necesario tener un equilibrio entre la protección de la privacidad y la seguridad y las oportunidades y beneficios que ofrece Internet.

1.2 Objetivos de la investigación.

Los objetivos de investigación que nos proponemos abordar relacionados con la anonimización, ocultación y eliminación de la huella digital generalmente se van a centrar en abordar cuestiones clave relacionadas con la privacidad, seguridad y protección de datos.

- En el ámbito de desarrollo de técnicas efectivas, investigaremos y comentaremos técnicas de anonimización que preserven la utilidad de los datos mientras protegen la identidad de los individuos. Además, exploraremos métodos de ocultación de información sensible de manera efectiva comparando el compromiso entre calidad y utilidad de los datos.
- Haremos una evaluación de riesgos y vulnerabilidades analizando los riesgos asociados con la identificación de individuos a través de huellas digitales, incluso cuando los datos han sido anonimizados, ya que habitualmente nos enfrentamos al riesgo de que se pueda revertir la anonimización, haciendo posible la reidentificación de personas. Razón por la cual, la anonimización no puede basarse solo en la aplicación de reglas determinadas, sino que también debe tener en cuenta las probabilidades de una posible reidentificación, con el objetivo de minimizar los riesgos para los derechos y libertades de los individuos.
- Identificaremos vulnerabilidades en diversas técnicas de anonimización existentes y propondremos soluciones para abordar estas debilidades.
- Discutiremos sobre cumplimiento normativo y comentaremos regulaciones y leyes de privacidad existentes en diferentes países, como el Reglamento General de Protección de Datos (GDPR) en la Unión Europea.
- Dentro de las herramientas automatizadas y buenas prácticas, comentaremos diferentes medidas y técnicas a adoptar que faciliten y proporcionen anonimización, ocultación y eliminación de huellas digitales.
- Mencionaremos la importancia de la educación y concientización dentro del mundo TI viendo como la fortaleza de la seguridad reside en el eslabón más débil de la cadena.
- Definiremos casos de uso específicos y como aplicar de técnicas de anonimización en entornos de atención médica, finanzas o como levantar y configurar soluciones determinadas que nos aseguren privacidad en entornos digitales.
- Analizaremos los desafíos y las limitaciones a las que nos enfrentamos para poder mantener el anonimato y la privacidad dentro del mundo de las comunicaciones digitales y la influencia que las nuevas tecnologías tienen sobre él.

1.3 Justificación del tema.

En un mundo cada vez más inmerso en la era digital, la preservación de la privacidad y la seguridad de los datos personales emerge como una temática de vital importancia. La "Anonimización, ocultación y eliminación de la huella digital" se erige como una disciplina esencial en el ámbito de la tecnología de la información, dirigida de manera específica a la salvaguarda de la privacidad de los usuarios en un entorno digital dinámico y expansivo.

A lo largo de este trabajo, nos embarcaremos en una exhaustiva revisión de las diversas técnicas relacionadas con la anonimización, ocultación y eliminación de la huella digital, explorando su aplicación en diferentes contextos y abordando los desafíos y limitaciones inherentes, así como las soluciones propuestas para superarlos.

En el ámbito de la anonimización, nos sumergiremos en un análisis detallado de múltiples técnicas y clasificaciones. Si bien existe una diversidad de enfoques, nos centraremos de manera específica en las técnicas de aleatorización, generalización y seudonimización, desentrañando sus aplicaciones prácticas.

El trabajo también arrojará luz sobre la ocultación y eliminación de la huella digital, brindando una revisión completa de las técnicas y destacando puntos esenciales para minimizar la huella digital, asegurando así la privacidad de los individuos mientras navegan por la vastedad de Internet.

Se nombrarán diferentes implicaciones normativas y legales de la anonimización, ocultación y eliminación de la huella digital en os diferentes países, nombrando las ideas más relevantes de estas normativas en términos de privacidad, protección de datos personales, acceso a la información y libertad de expresión.

Este trabajo busca proporcionar una comprensión profunda de las técnicas de anonimización y ocultación y contextualizarlas en el marco de las leyes y regulaciones vigentes en diferentes jurisdicciones. En un mundo digital en constante evolución, el análisis riguroso de estas prácticas se revela esencial para promover una convivencia armoniosa entre la tecnología y la salvaguarda de los derechos individuales.

2 FUNDAMENTOS TEÓRICOS

2.1 Conceptos básicos de anonimización y protección de datos.

2.1.1 Introducción.

La protección de datos se considera un tema de suma importancia en la sociedad actual por diversas razones, entre ellas el enorme desarrollo que ha experimentado la era digital.

Por un lado, se considera un derecho fundamental reconocido en los artículos 18.4 de la Constitución Española [1], 8.1 de la Carta de los Derechos Fundamentales de la Unión Europea [2] y el 16.1 del Tratado de Funcionamiento de la Unión Europa [3]. Este derecho garantiza a la persona el control sobre sus datos, tanto personales como no personales, y sobre su uso y destino, para evitar el tráfico ilícito de los mismos o lesivo para la dignidad y los derechos de los afectados [4].

Por otro lado, la protección de datos mejora y refuerza la seguridad de la información, lo que es vital para prevenir el uso indebido de los datos y la aparición de ciberdelincuentes [5].

Por último, mencionar que, en muchos países, la protección de datos es un requisito legal. Por ejemplo, en la Unión Europea, la protección de datos es un derecho fundamental y el Reglamento General de Protección de Datos (RGPD) es el marco regulatorio para este derecho [6].

2.1.2 Privacidad digital.

La privacidad era un tema delicado desde mucho antes de la aparición de las computadoras. Sin embargo, las preocupaciones han aumentado debido a la existencia y al uso generalizado de grandes bases de datos que facilitan la compilación de información sobre un individuo a partir de muchas fuentes de datos diferentes. Los problemas de privacidad se ven aún más exacerbados ahora que Internet hace que sea más fácil la recogida automática de nuevos datos y se añadan a las bases de datos [7]. Hoy en día, los datos introducidos en formularios o contenidos en bases de datos existentes pueden combinarse casi sin esfuerzo. A medida que las herramientas y servicios de minería de datos se vuelvan más accesibles, es probable que las preocupaciones sobre la privacidad aumenten aún más [8].

2.2 Definición de huella digital.

Huella digital o huella electrónica, en este contexto, hace referencia a la serie de datos que se generan al utilizar Internet. Esto abarca desde los sitios web que se visitan hasta los correos electrónicos y la información que se comparte en línea. Esta huella digital puede ser empleada para seguir las actividades y los dispositivos en línea de una persona [9].

Cada vez que se interactúa en línea, ya sea para publicar en redes sociales, suscribirse a un boletín informativo, dejar una reseña o comprar en línea, se amplía la huella digital. Pero no solo eso, también

hay formas menos evidentes de dejar rastro en la red. Por ejemplo, los sitios web pueden usar cookies para seguir actividades en línea, y las aplicaciones pueden recoger datos sin consentimiento. Además, la información personal puede ser vendida o compartida con otras organizaciones, o incluso filtrada por hackers.

De manera genérica la huella digital se genera a partir de los siguientes tipos de datos [10]:

- **Datos públicos:** son los datos relativos a declaraciones de impuestos, direcciones en las facturas de servicios, resúmenes de tarjetas de crédito, detalles de cargos, datos relacionados con becas, resultados de sorteos, resoluciones judiciales, etc.
- **Datos publicados por otros:** son publicaciones o fotos que realizan amigos, familiares, asociaciones o clubs, perfiles públicos en redes sociales, fotos, etc.
- **Datos que genera uno mismo:** son las interacciones que se realizan en redes sociales como publicaciones, reenvíos, likes, comentarios, fotos y foros. Formularios completados, contenidos compartidos en diferentes plataformas como perfiles en redes de contactos, currículum vitae u otros contenidos diversos como listas de reproducción y videos favoritos.

2.2.1 Tipos de huella digital y su importancia

Hay dos categorías de huellas digitales: las activas y las pasivas, que se distinguen por el consentimiento informado. La huella digital activa se forma a través de las actividades de intercambio de datos en línea que realiza intencionadamente o con su consentimiento informado. En contraste, la huella digital pasiva se crea a partir de los datos recopilados cuando sus actividades en línea son rastreadas sin su consentimiento informado ni su conocimiento.

Huella digital activa

La huella digital activa es el rastro que se deja en línea cuando se comparte información sobre las personas de forma deliberada. Esto incluye publicaciones en redes sociales, foros y aplicaciones que solicitan permisos, como la ubicación o datos personales [11].

Este tipo de huella, puede generarse en diversas situaciones, pero su característica fundamental se basa en que el usuario que dejó la información es conocido.

En un entorno en línea, la huella digital activa puede almacenarse tan pronto como el usuario se conecte y se autentique en cualquier sitio web de la Red. Un claro ejemplo de esto sería el acceso a una red social para subir contenido, o cuando escribe un mensaje en un foro o edita una entrada en un blog específico.

En un entorno cerrado u *off-line*, es decir, fuera de línea, el rastro dejado por el usuario podría ser almacenado en archivos temporales, similar a lo que ocurre con la huella digital pasiva. Sin embargo, existe una diferencia crucial: el administrador tiene un control sobre las cuentas asociadas a usuarios específicos [12].

Huella digital pasiva

Definimos como huella digital pasiva toda la información que proporcionamos y que es extraída de forma inconsciente, a veces con nuestro consentimiento y otras veces de manera más cuestionable. En cualquier caso, se trata de una huella digital pasiva, ya que el usuario no comparte activamente información personal, sino de forma implícita. Un ejemplo destacado de huella digital pasiva es la recopilación de datos a través de las cookies, que determinan aspectos de los visitantes en páginas web, como su ubicación, el tipo de navegador que utilizan, el tiempo que pasan en el sitio web y las páginas que visitan.

En el contexto de Internet, que es una red global, la huella digital puede ser almacenada en línea como una entrada en la base de datos de cualquier servidor de servicios al que hagamos una solicitud. Esta huella puede consistir en la dirección IP que el usuario utiliza para conectarse, junto con otros

detalles como el origen de esa IP y la fecha en que fue creada. La información almacenada puede ser revisada posteriormente mediante herramientas de análisis de red, si alguien lo considera necesario.

En un ámbito más local, como una red de área local, la huella digital puede ser registrada en archivos temporales en forma de registro (log). Solo el administrador de la red tendría acceso a esta información, que podría incluir datos sobre las acciones realizadas desde una máquina específica. El inconveniente en este caso es que, si el administrador no tiene un sistema de control de cuentas, puede resultar imposible determinar qué usuario llevó a cabo una acción específica [12].

La importancia de las huellas digitales reside en las siguientes razones [13]:

- La **permanencia** relativa, una vez que los datos se vuelven públicos (o incluso semipúblicos, ej: las publicaciones en Facebook), el propietario tiene escaso control sobre cómo serán utilizados por otros.
- La **reputación digital:** una huella digital puede determinar la reputación digital de una persona. Actualmente dicha reputación es casi tan importante como la reputación fuera de Internet.
- **Verificación de huella digital:** los empleadores tienen la capacidad de revisar las huellas digitales de los candidatos a empleo, especialmente sus perfiles en redes sociales, antes de tomar decisiones de contratación. De manera similar, las instituciones educativas, como colegios y universidades, pueden examinar las huellas digitales de sus posibles estudiantes antes de admitirlos.
- **Malinterpretación o modificación de contenido:** las palabras y las imágenes que se comparten en línea pueden ser malinterpretadas o alteradas, lo que podría dar lugar a ofensas involuntarias.
- **Difusión de contenido:** el contenido destinado a un grupo privado puede difundirse a un círculo más amplio sin permiso ni control, con el riesgo de daño de relaciones y amistades.
- **Cibercrimen**: los delincuentes cibernéticos pueden aprovechar la huella digital, utilizando esta información para llevar a cabo actividades como el phishing, accediendo a cuentas o creando identidades falsas basadas en los datos recopilados.

Los expertos en seguridad informática afirman que los smartphones son nuestra principal fuente de rastros digitales. En la actualidad, los usuarios almacenan en sus teléfonos inteligentes no solo fotos personales, listas de contactos y archivos de trabajo, sino también correos electrónicos, contraseñas e incluso aplicaciones para interactuar con el banco y otras empresas con las que contratan servicios. Un simple robo de nuestro dispositivo permitiría acceder a toda esta información.

A pesar de que los gestores de las principales redes sociales insisten en que los usuarios son libres de establecer el nivel de privacidad que deseen y que pueden eliminar los datos personales que prefieran, en Internet todo está interconectado y sitios como Google o Archive.org registran la memoria completa de la red. Por lo tanto, independientemente de las políticas de privacidad que de adopten, cualquier dato personal que se haya subido alguna vez a la red quedará registrado de alguna manera.

2.2.2 Importancia de la privacidad digital

La privacidad digital es esencial por múltiples razones. En primer lugar, nos permite mantener el control sobre nuestra información personal y decidir quién tiene acceso a ella. Esto es de vital importancia para prevenir el robo de identidad, el fraude y otras formas de abuso en línea. Por otro lado, la privacidad digital es crucial para preservar nuestra autonomía y libertad en el entorno digital. Al proteger nuestra información personal, se evita que terceros la utilicen de manera no autorizada o la compartan con fines publicitarios, políticos u otros propósitos que puedan comprometer nuestra privacidad y seguridad [14].

La privacidad digital es fundamental para nuestra seguridad, tanto en el ámbito digital como en el físico. Aunque en ciertas ocasiones sea necesario compartir información personal con terceros para llevar a cabo trámites o acceder a servicios específicos, es crucial gestionar estos datos de manera adecuada y protegerlos según las leyes establecidas. Si no se manejan correctamente, estos datos pueden ser objeto de uso indebido y abusos, con consecuencias diversas para nuestros derechos, libertades y, como ya mencionamos, nuestra seguridad.

En ocasiones, podemos subestimar el impacto de proporcionar algunos de nuestros datos personales a empresas o compañías, pensando que no afectará significativamente nuestras vidas. Sin embargo, solemos compartir mucha más información de la que imaginamos y lo hacemos sin comprender completamente cómo se utilizará esa información ni quién la utilizará. Esto sucede porque muchas veces no nos molestamos en leer las políticas de privacidad o de cookies, ni en configurar los ajustes de privacidad de nuestras cuentas de usuario y redes sociales.

Todos estos datos personales se acumulan, se cruzan y se analizan para inferir nueva información que puede tener repercusiones en nuestra vida diaria. Por ejemplo, si se deduce de nuestra información que tenemos problemas de salud, es posible que una compañía de seguros médicos o de vida nos niegue un contrato.

La protección de la privacidad digital es esencial porque nuestros datos e información personal pueden ser utilizados con diversos fines, algunos de los cuales no son éticos ni bien intencionados. Desde mostrarnos publicidad personalizada hasta manipular nuestras opiniones, nuestros datos pueden ser empleados para cometer diversos delitos, como la suplantación de identidad. La famosa frase "la información es poder" se aplica también a los datos personales, ya que son sumamente valiosos para empresas, entidades públicas y otros actores, incluyendo creadores de desinformación y delincuentes cibernéticos. Proteger nuestra privacidad en línea significa resguardarnos contra abusos, manipulaciones y posibles ciberdelitos.

2.2.3 Marco legal de la privacidad digital y regulaciones relacionadas.

En el ámbito de la privacidad digital, existen varios marcos legales y regulaciones en diferentes países y regiones del mundo para proteger los datos personales de los individuos.

Las regulaciones relacionadas con la privacidad digital se refieren al derecho de las personas a controlar sus datos personales, a proteger su identidad y su intimidad en el ámbito digital. Así, la privacidad digital implica cuestiones como el consentimiento, la transparencia, la seguridad y la responsabilidad de los datos.

Por estos motivos y por su importancia a nivel individual, la protección de la privacidad digital también es crucial para preservar elementos fundamentales como la integridad de los sistemas democráticos, la justicia, el cumplimiento de la ley y la libertad individual y colectiva [15].

Esta necesidad de proteger la privacidad digital ha llevado a la implementación de leyes específicas.

En **España**, dentro del marco legal de la privacidad digital podemos encontrar

- *Ley Orgánica 3/2018 de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD)*, que establecen obligaciones claras en cuanto a la protección de datos personales para organizaciones y profesionales.
- *Ley de Servicios de la Sociedad de la Información y de Comercio Electrónico (LSSI)*: La LSSI es una ley española que regula el uso de las tecnologías de la información y la comunicación (TIC) para la prestación de servicios de la sociedad de la información. La LSSI establece una serie de requisitos para las organizaciones que recopilan o utilizan datos personales en el contexto de la prestación de servicios de la sociedad de la información.

En la **Unión Europea**, el marco legal más relevante es el *Reglamento General de Protección de Datos (RGPD),* que entró en vigor en 2018 y establece las normas comunes para el tratamiento de los datos personales por parte de las autoridades públicas, las empresas y otras organizaciones. El RGPD otorga a los ciudadanos de la UE derechos como el de acceso, rectificación, supresión, limitación, portabilidad y oposición de sus datos, así como el de presentar reclamaciones ante las autoridades de protección de datos. El RGPD se aplica a todas las organizaciones que traten datos personales de residentes de la UE, independientemente de su ubicación [15].

Además del RGPD, la UE cuenta con otras normas relacionadas con la privacidad digital, como la *Directiva sobre protección de datos en el ámbito penal*, que regula el intercambio de datos personales entre las autoridades policiales y judiciales de los Estados miembros, y la *Directiva sobre la privacidad y las comunicaciones electrónicas*, que regula aspectos como las cookies, el spam y la confidencialidad de las comunicaciones.

Fuera de la UE, existen otras leyes y normativas de Internet que afectan a la privacidad digital, como la *Ley de Protección de la Privacidad en Línea para Niños (COPPA)* de Estados Unidos, que protege la información personal de los menores de 13 años, o *la Ley de Privacidad del Consumidor de California (CCPA),* esta ley, vigente desde enero de 2020, otorga a los residentes de California el derecho a saber qué datos personales se recopilan sobre ellos, así como el derecho a optar por no compartir su información y solicitar la eliminación de datos. La *Ley de Privacidad del Consumidor de Virginia (VCDPA),* esta ley, en vigor desde marzo de 2021, otorga a los residentes de Virginia derechos similares a los de la CCPA, permitiéndoles controlar su información personal y saber cómo las empresas la utilizan. La *Ley de Privacidad del Consumidor de Colorado (CPRA)*, similar a la CCPA de California, esta ley establece derechos de privacidad para los residentes de Colorado, permitiendo a las personas acceder y controlar la información personal que las empresas recopilan sobre ellos. La *Ley de Protección de Datos Personales en Brasil (LGPD),* esta ley, en vigor desde septiembre de 2020, establece normas para el tratamiento de datos personales en Brasil, otorgando a los individuos derechos sobre su información personal y estableciendo obligaciones para las organizaciones que procesan estos datos. La *Ley de Protección de Información Personal en Japón*, esta ley regula el uso y la protección de la información personal en Japón, imponiendo restricciones sobre la recopilación y el manejo de datos personales.

Incluso en la ONU, tanto Asamblea General de las Naciones Unidas y el Consejo de Derechos Humanos han aprobado numerosas resoluciones sobre el derecho a la privacidad en la era digital. Resolución de la AG 75/176, Resolución del CDH 42/15, Resolución de la AG 73/179, Resolución del CDH 37/2, Resolución del CDH 34/7, Resolución de la AG 71/199, Resolución del CDH 28/16, Resolución de la AG 69/166, Resolución de la AG 68/167 [16].

Estas resoluciones subrayan la responsabilidad de los Estados de asegurar que cualquier intervención en el derecho a la privacidad esté en consonancia con los principios de legalidad, necesidad y proporcionalidad. Además, establece que los derechos fundamentales de las personas, incluyendo el derecho a la privacidad, deben ser protegidos también en el ámbito de Internet. Reconoce asimismo que la adopción, implementación y avance de nuevas tecnologías emergentes, como la inteligencia artificial, pueden impactar en el ejercicio del derecho a la privacidad y otros derechos humanos.

3 ANONIMIZACIÓN DE DATOS

3.1 Definición y objetivos

La anonimización de datos se refiere a la metodología y conjunto de buenas prácticas y técnicas diseñadas para reducir el riesgo de identificación de personas. Este proceso es esencial para salvaguardar la privacidad de los individuos, asegurando el cumplimiento de las normativas y derechos fundamentales[17].

Los objetivos de la anonimización de datos son los siguientes:

- Minimizar el riesgo de identificación de personas.
- Garantizar la irreversibilidad del proceso de anonimización.
- Realizar una auditoría de la utilización de los datos anonimizados, supervisando quién los usa, cuándo y con qué propósito.

Además, la anonimización de datos busca preservar la integridad de la información almacenada o compartida y asegurar el cumplimiento de rigurosas regulaciones de privacidad de datos. Según el estándar ISO (ISO 29100:2011), el principal criterio de anonimización consiste en modificar de manera irreversible la información personal identificable (PII) para que la identidad de la persona ya no pueda ser determinada directa o indirectamente [18].

3.2 Técnicas de anonimización

Las diferentes técnicas conocidas y desarrolladas para conseguir la anonimización de los datos combinan metodología, buenas prácticas y técnicas para proteger la privacidad de las personas, garantizar el cumplimiento normativo y permitir el uso responsable de los datos.

Dentro de las diferentes técnicas de anonimización vamos a enfocar nuestro estudio en la descripción, por un lado, de propuestas teóricas y por otro lado de técnicas o soluciones prácticas que se utilizan ampliamente hoy en día.

3.2.1 Propuestas teóricas

Dentro del marco teórico de estudio concerniente a los diferentes enfoques de anonimización, podemos definir tres enfoques generales:

- **Aleatorización**: Dentro de este enfoque, se elimina la correlación entre los datos y el individuo, ayudándonos de la adición de ruido, la permutación, o la Privacidad Diferencial [19].

- **Generalización**: alteración de escalas u órdenes de magnitud a través de técnicas basadas en agregación como Anonimato-K, Diversidad-L, Proximidad-T o Reducción de la precisión de los datos.
- **Seudonimización o pseudonimato**: es un proceso en el cual se reemplaza la información identificable de una persona con un seudónimo o identificador único, de tal manera que se hace más difícil o imposible vincular los datos al individuo original sin información adicional. De esta manera, se reemplazan los valores por versiones cifradas o tokens, habitualmente mediante algoritmos de HASH, impidiendo la identificación directa del individuo, a menos que se combine con otros datos adicionales, que deben estar custodiados de forma adecuada.

3.2.2 Aleatorización

La aleatorización se refiere a la asignación en grupos de manera impredecible de los objetos que participen en un estudio. El propósito principal de la aleatorización es asegurar que los grupos involucrados en el experimento estén equilibrados, de modo que sean similares en cuanto a la distribución de todos los factores, ya sean conocidos o desconocidos, que podrían influir en los resultados del estudio. Esto significa que cualquier diferencia entre los dos grupos probablemente se deba al efecto de la intervención que se está estudiando [20].

Adición de ruido

La aleatorización es un conjunto de técnicas utilizadas para alterar la precisión de los datos con el propósito de eliminar cualquier conexión identificable entre los datos y sus propietarios. En el ámbito de la aleatorización, se encuentra la técnica conocida como adición de ruido.

La adición de ruido implica la modificación de los atributos de un conjunto de datos de manera que se vuelvan menos precisos, manteniendo, sin embargo, su distribución general. Cuando esta técnica se aplica de manera adecuada, un tercero no debería ser capaz de identificar a una persona ni de recuperar los datos originales ni de descifrar cómo han sido alterados.

Por lo general, la adición de ruido se combina con otras técnicas de anonimización, como la eliminación de atributos evidentes y de cuasi identificadores. El nivel de ruido que se agrega depende de la cantidad y naturaleza de la información requerida, así como del impacto que la revelación de los atributos protegidos pueda tener en la privacidad de las personas.

Es fundamental comprender que la adición de ruido es una medida complementaria destinada a dificultar que un posible atacante acceda a los datos personales. No obstante, no se debe considerar una solución completa para la anonimización, a menos que el nivel de ruido supere la información contenida en el conjunto de datos [21].

Permutación

La aleatorización de bloques permutados es un método para asignar de manera aleatoria a los participantes a diferentes grupos de tratamiento, al mismo tiempo que se asegura que exista un equilibrio entre estos grupos. Cada "bloque" contiene un número predeterminado de asignaciones de tratamiento que se organizan de forma aleatoria [22].

Aleatorización estratificada

Este modelo se asemeja al de los bloques, pero segmenta los grupos en varios subgrupos o estratos. Esta segmentación se realiza considerando un factor relevante que se cree que puede afectar los resultados finales. La división se realiza según puntos de corte, que generalmente se basan en la información obtenida de estudios anteriores [23].

Aleatorización mediante minimización (aleatorización adaptativa)

La aleatorización mediante minimización o aleatorización adaptativa es un método de asignar individuos a diferentes grupos con el objetivo de minimizar las diferencias entre estos grupos, no solo en cuanto al número de individuos, sino también en cuanto a las características de los mismos [24].

Este método se realiza con tamaños reducidos de muestra [25].

Privacidad diferencial

La privacidad diferencial comprende un conjunto de técnicas que nos posibilitan recopilar y compartir datos con la "certeza matemática" de que aquellos individuos que suministraron dichos datos no serán de ninguna manera perjudicados o identificados [26]. No es un algoritmo, sino un sistema o marco de trabajo para mejorar la privacidad de los datos [27].

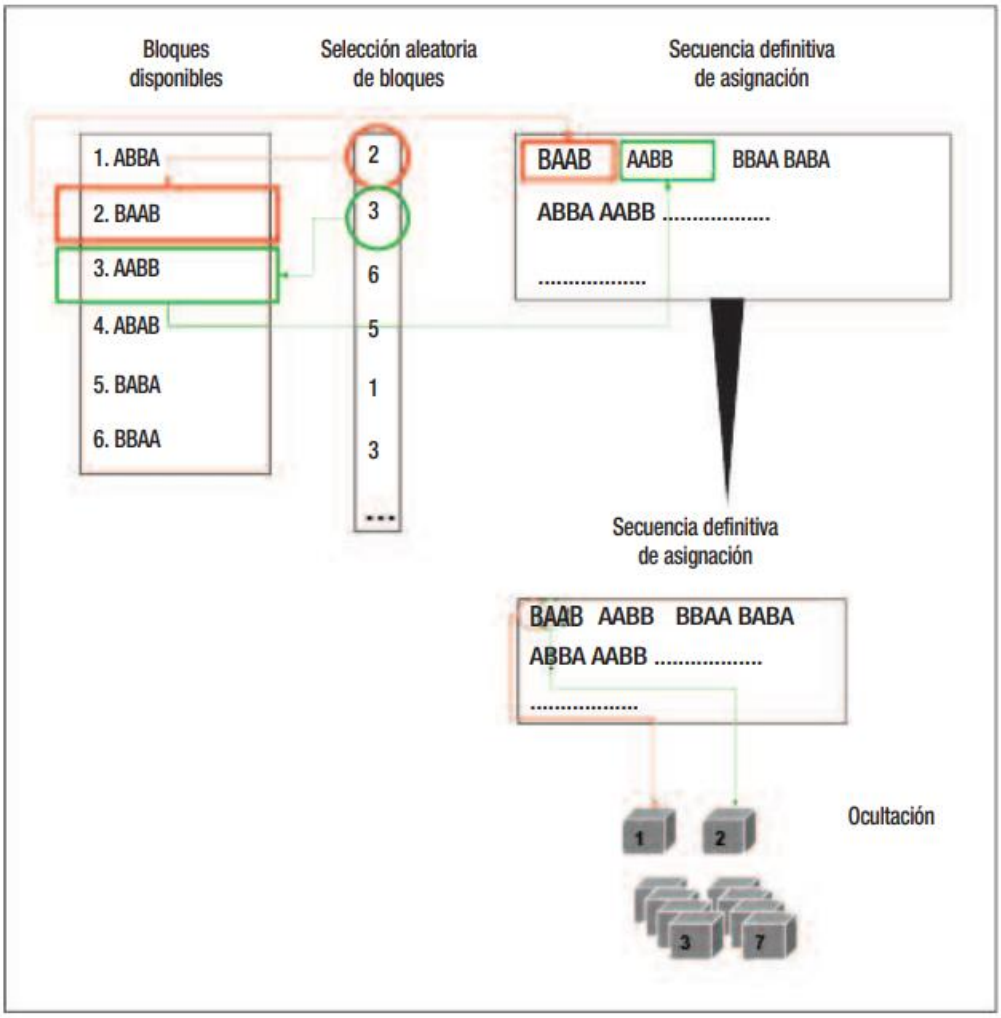


Figura 3-1 Ejemplo de funcionamiento de la privacidad diferencial [28]

Simplificando, consideramos que un sistema es diferencialmente privado cuando los datos están estructurados de tal manera que no podemos conocer si un sujeto en concreto participó o no [29]. Cuando un sistema satisface esta condición, los datos no pueden ser asociados a las personas, preservando así su privacidad.

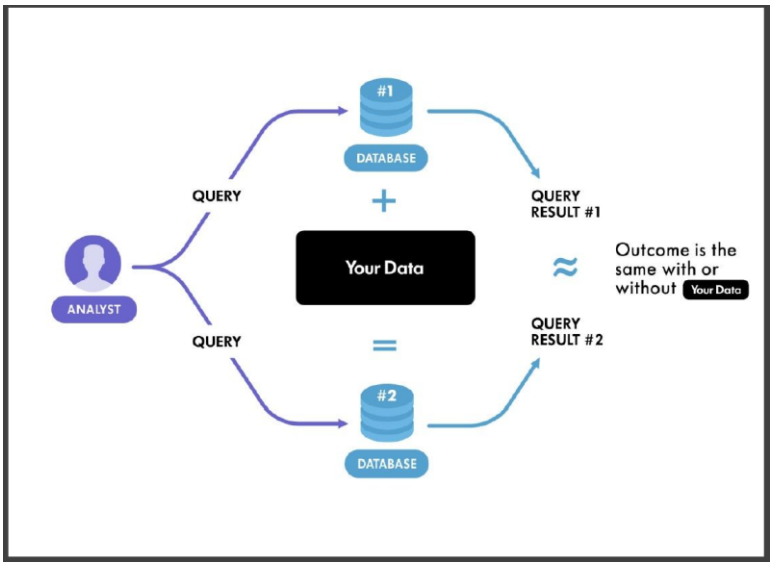


Figura 3-2 Anonimato proporcionado por la privacidad diferencial [30]

- **Privacidad Diferencial Epsilon:** es un concepto utilizado para medir la cantidad de información que se divulga acerca de un individuo al realizar una consulta en un conjunto de datos. A medida que el valor de Epsilon disminuye, aumenta el nivel de privacidad, ya que se revela menos información sobre el individuo en cuestión [31].

  Es decir, la cantidad que cada individuo contribuye al resultado de una consulta en una base de datos depende en parte de cuántos datos de personas están involucrados en la consulta. Si la base de datos contiene datos de una sola persona, los datos de esa persona contribuyen con el 100%. Si la base de datos contiene datos de cien personas, los datos de cada persona contribuyen solo con un 1%.

- **Privacidad Diferencial Local (Local Differential Privacy - LDP):** En el enfoque de privacidad diferencial local, cada individuo que contribuye con datos agrega ruido o perturbación a sus propios datos antes de enviarlos al servidor o realizar análisis locales [32]. Cada individuo controla su propia privacidad y contribuye a la protección de sus datos personales. Luego, en el servidor o en el análisis centralizado, se aplica una agregación de estos datos ruidosos para obtener estadísticas o resultados globales. La privacidad se garantiza a nivel de cada contribuyente individual [31].

- **Privacidad Diferencial Global (Global Differential Privacy - GDP)**: En este enfoque, se garantiza la privacidad de manera global para todo el conjunto de datos. Los datos en crudo de las personas son recopilados y analizados por un organismo central, aplicando los algoritmos de privacidad diferencial al conjunto de datos [33]. Se aplica una cantidad controlada de ruido o perturbación a nivel del conjunto completo de datos antes de que se realicen consultas o análisis. El nivel de ruido se ajusta de tal manera que protege la privacidad de todas las personas en el conjunto de datos. Esto significa que, incluso si alguien tiene un conocimiento detallado sobre la mayoría de los datos, aún es extremadamente difícil identificar información sobre individuos específicos.

- **Privacidad Diferencial por Diseño (PbD):** Es una estrategia preventiva para salvaguardar la privacidad de los datos que se ideó en la década de 1990 [34]. En sus inicios, su objetivo principal era crear un modelo sólido y adaptable para proteger la privacidad de los datos, superando así las "tecnologías de mejora de la privacidad" y los estándares de cumplimiento normativo menos robustos que existían en ese momento. Todo esto con la finalidad de garantizar una privacidad de datos completa [35].

- **Privacidad Diferencial Compuesta:** Este enfoque se emplea para salvaguardar la privacidad de los datos personales en situaciones donde se realizan consultas que involucran múltiples operaciones en los datos [36].

3.2.3 Generalización

La técnica de anonimización conocida como "generalización" es una estrategia fundamental en la protección de la privacidad de los datos personales. Su principal objetivo es reducir la precisión de la información identificable de manera intencionada, pero preservando la integridad del conjunto de datos. Esto se logra al modificar los datos de manera que se vuelvan menos específicos, lo que dificulta la tarea de vincular esos datos con individuos particulares.

Un ejemplo ilustrativo de la generalización se encuentra en la gestión de edades en una base de datos. En lugar de almacenar las edades exactas de las personas, se aplicaría la generalización reemplazando las edades precisas con rangos de edades. Por ejemplo, en lugar de registrar que alguien tiene 32 años, se podría indicar que la persona se encuentra en el rango de 30-35 años. Esta práctica mantiene la utilidad general de los datos para análisis estadísticos y tendencias, pero hace más complicado el proceso de identificar a una persona específica basándose únicamente en su edad.

La importancia de la generalización radica en su capacidad para salvaguardar la privacidad de los individuos sin comprometer la utilidad de los datos para análisis legítimos. Sin embargo, es fundamental aplicar la generalización con cautela, ya que un exceso de generalización puede llevar a que los datos sean demasiado vagos y, por lo tanto, inútiles para cualquier análisis posterior.

K-Anonimato

El K-anonimato es un modelo de protección formal de la información que se aplica junto a un conjunto de políticas de seguridad que proporcionan protección adicional. Una publicación de datos proporciona protección de k-anonimato si la información de cada individuo contenida en la publicación no puede ser distinguida de al menos k-1 individuos cuya información también aparece en la publicación [37].

L-Diversidad

Se ha expuesto que publicar datos sobre individuos sin revelar información sensible sobre ellos es un problema importante. Con la aparición del k-anonimato se desarrolló una técnica para mejorar la privacidad.

Sin embargo, se ha demostrado que mediante dos simples ataques un conjunto de datos k-anonimizado presenta algunos problemas de privacidad. Primero, un atacante puede descubrir valores de atributos sensibles cuando hay poca diversidad en los mismos. Este es un problema conocido. En segundo lugar, los atacantes a menudo tienen conocimientos previos, de tal manera que el k-anonimato no garantiza la privacidad contra los atacantes [38].

Para resolver las limitaciones de modelo K-anonimato, se ha desarrollado un nuevo modelo de privacidad llamado L-diversidad que mejora el anonimato ya que puede defenderse contra estos ataques [39].

La L-diversidad se basa en la idea de que, en un conjunto de datos, cada registro (o fila) debe tener al menos "L" valores distintos en ciertos atributos sensibles. Estos atributos sensibles son aquellos que podrían utilizarse para identificar o inferir información personal sobre los individuos en el conjunto de datos. El valor "L" se refiere al número mínimo de valores diferentes que deben existir en estos atributos para garantizar la privacidad [40].

Por ejemplo, si estamos trabajando con un conjunto de datos de pacientes médicos y uno de los atributos es la edad, podríamos aplicar la L-diversidad estableciendo que cada grupo de registros con la misma edad debe tener al menos "L" pacientes diferentes. Esto dificulta la identificación de un

individuo específico en función de su edad, ya que hay varias personas con la misma edad en ese grupo.

El objetivo de la L-diversidad es prevenir ataques de reidentificación, donde un atacante podría utilizar información en el conjunto de datos para identificar a personas específicas. Al garantizar que haya suficiente diversidad en los atributos sensibles, se reduce el riesgo de que estos ataques tengan éxito.

T-Proximidad

La T-Proximidad es un concepto relacionado con la protección de la privacidad de los datos en conjuntos de datos sensibles. Al igual que la L-diversidad, la T-proximidad es una técnica utilizada para mitigar los riesgos de la reidentificación de individuos a partir de datos publicados. Su objetivo principal es garantizar que las distribuciones de ciertos atributos sensibles en un conjunto de datos sean similares a las distribuciones en la población general, lo que hace que sea más difícil identificar a personas específicas.

La T-proximidad se basa en la noción de distancia estadística entre las distribuciones de los valores de un atributo sensible en un conjunto de datos y en la población general. En otras palabras, se busca medir cuán cercanas son estas distribuciones para evitar que un atacante pueda identificar a una persona en el conjunto de datos en función de un atributo sensible [41].

Para lograr la T-proximidad, se establece un parámetro "T" que representa la distancia máxima permitida entre las distribuciones. Si la distancia entre las distribuciones es menor o igual a "T", se considera que el conjunto de datos cumple con la T-proximidad.

Por ejemplo, si estamos trabajando con un conjunto de datos médicos y uno de los atributos sensibles es la presión arterial, la T-proximidad aseguraría que la distribución de los valores de presión arterial en el conjunto de datos sea similar a la distribución de presión arterial en la población general. Esto dificulta que un atacante pueda identificar a una persona en función de su presión arterial, ya que los valores en el conjunto de datos son representativos de la población en general.

Reducción de la precisión de los datos

La reducción de la precisión de datos consiste en disminuir el nivel de detalle de los datos, de forma que se pierda información que pueda identificar a un objeto. Se puede reducir la precisión geoespacial redondeando las coordenadas geográficas a un nivel de precisión menor, la precisión temporal, pasando de día/mes/año a solo año, o la precisión numérica agrupando valores en rangos de valores. La reducción de la precisión de datos puede hacerse mediante técnicas como el redondeo, el truncamiento o el agrupamiento.

La reducción de la precisión de datos tiene como ventaja que es una técnica simple y fácil de aplicar, que puede reducir el tamaño de los datos y facilitar su análisis. Sin embargo, también tiene como inconveniente que puede provocar una pérdida de información relevante, que puede afectar a la calidad y la utilidad de los datos. Además, la reducción de la precisión de datos no garantiza que los datos sean totalmente anónimos, ya que puede haber otros atributos que permitan la reidentificación de las personas

3.2.4 Seudonimización o pseudonimato

La seudonimización o pseudonimato consiste en reemplazar mediante la utilización de alguna técnica la información que puede ser considerada sensible, dentro de esta información podemos encontrar la PII, PHI y PCI.

La información de identificación personal (PII) que permiten identificar a una persona física por otros datos que no lo permiten, pero que mantienen una relación con los originales mediante una información adicional.

La información de salud personal (PHI) que abarca una amplia gama de datos. La PHI está definida por la Ley de Portabilidad y Responsabilidad de los Seguros Médicos (HIPAA) y comprende cualquier información que pueda ser utilizada para vincular la identidad de una persona con su historial médico [42].

La información de la Industria de Tarjetas de Pago (PCI, por sus siglas en inglés) es cualquier dato utilizado durante una transacción con tarjeta de pago y se superpone para incluir información personal identificable (PII), por lo que sí, la PCI abarca la PII. Este tipo de datos está generalmente asociado con el sector de servicios financieros. Debido a los cambios continuos en la gobernanza de la información PCI según los requisitos del Estándar de Seguridad de Datos de la Industria de Tarjetas de Pago (PCI-DSS), todas las organizaciones que aceptan o procesan tarjetas de crédito como método de pago deben estar informadas sobre los requisitos para salvaguardar la información PCI [43].

Característica	PII	PCI	PHI
Definición	Datos que pueden identificar a una persona específica.	Datos utilizados en transacciones con tarjetas de pago.	Datos relacionados con la salud de una persona, identificables y protegidos por leyes específicas.
Ejemplos	Nombre, dirección, número de seguro social, etc.	Número de tarjeta de crédito, fecha de vencimiento, código de seguridad, etc.	Historial médico, información sobre tratamientos, identificadores de salud, etc.
Ámbito de aplicación	Amplio, abarca cualquier información que pueda identificar a una persona.	Específico para datos relacionados con transacciones de tarjetas de pago.	Relacionado con la atención médica y la salud de los individuos.
Regulación principal	Leyes de privacidad y protección de datos.	Estándar de Seguridad de Datos de la Industria de Tarjetas de Pago (PCI DSS).	Ley de Portabilidad y Responsabilidad de los Seguros Médicos (HIPAA) en los Estados Unidos, y regulaciones de privacidad en otras regiones.
Sectores de aplicación típicos	Cualquier sector que maneje datos personales.	Comercio minorista, servicios financieros.	Instituciones de atención médica, proveedores de servicios de salud.

Tabla 3-1 Tabla comparativa PII, PCI, PHI (Elaboración propia)

A diferencia de la anonimización tradicional, donde los datos se alteran de tal manera que se vuelven irreversibles y no pueden vincularse con el individuo original, la seudonimización permite almacenar información adicional que puede ser utilizada para volver a identificar a la persona si es necesario, pero solo por personas autorizadas. Esta información adicional se almacena por separado y se protege con medidas técnicas y organizativas para evitar la reidentificación de los interesados. La seudonimización es una forma de proteger la privacidad de los datos personales y reducir los riesgos de su tratamiento.

Como hemos comentado, la seudonimización es un proceso reversible, mientras que la anonimización no lo es.

Otra distinción fundamental entre anonimización y seudonimización radica en las medidas que aseguran la salvaguardia de los derechos de las personas afectadas. Los datos anonimizados, al dejar de ser considerados información personal, quedan fuera del alcance del RGPD, a diferencia de los

datos seudonimizados y la información adicional asociada a ellos, que sí están sujetos a las regulaciones del RGPD [44].

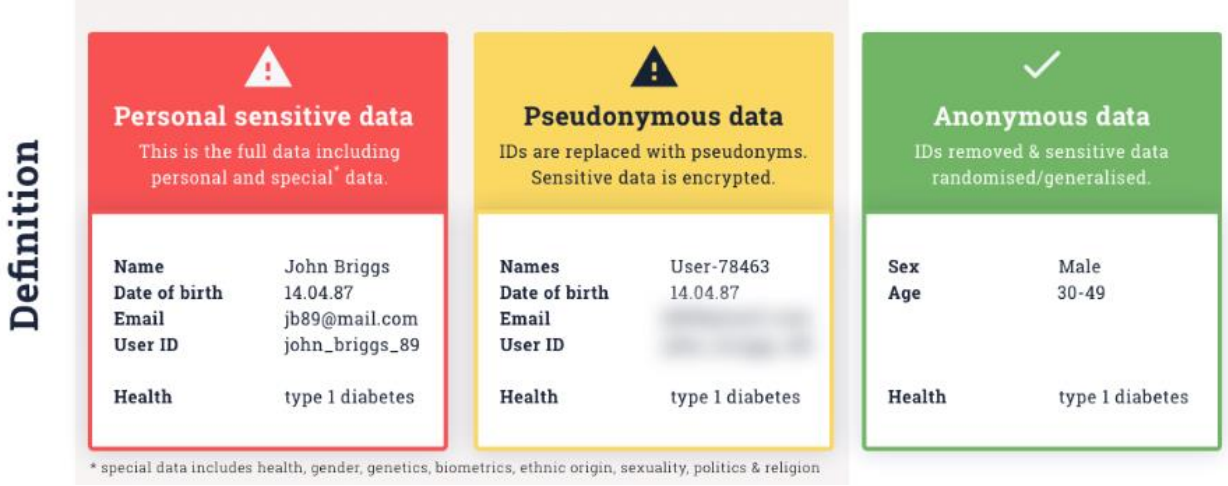


Figura 3-3 Diferencia entre anonimización y seudonimización [45].

En el proceso de seudonimización, es fundamental asegurar que los dos conjuntos de datos involucrados, *los datos seudonimizados y la información adicional asociada a ellos*, estén respaldados por las siguientes salvaguardas [44]:

- **Prevención de Reidentificación:** El procedimiento de seudonimización debe ser lo suficientemente sólido como para evitar la reidentificación de individuos si no se cuenta con la información adicional correspondiente.
- **Cumplimiento de Limitaciones del RGPD:** Se deben respetar todas las restricciones establecidas por el Reglamento General de Protección de Datos (RGPD) con respecto a las finalidades de la seudonimización, el período de retención de datos seudonimizados y la comunicación de dichos datos, entre otras restricciones.
- **Garantías Adicionales según el Riesgo:** Las medidas de seudonimización deben ser adaptadas de acuerdo al nivel de riesgo para los derechos y libertades de las personas físicas, incorporando garantías adicionales en consecuencia.
- **Garantías Técnicas y Organizativas:** Es esencial implementar garantías técnicas y organizativas robustas para prevenir cualquier violación de datos personales. Estas medidas deben aplicarse tanto al conjunto de datos seudonimizado como a la información adicional asociada, con el objetivo de evitar cualquier brecha de seguridad que pueda comprometer la privacidad de las personas involucradas.

Entre las técnicas más comunes de seudonimización podemos encontrar la *seudonimización de un único identificador*

3.2.4.1 Seudonimización de un Único Identificador

Dentro de la seudonimización de un único identificador, se analizan una lista de posibles enfoques, junto con las ventajas y limitaciones pertinentes.

Contador

El contador es la función de seudonimización más simple. Los identificadores son sustituidos por un número elegido por un contador monótono. En primer lugar, se establece una semilla s en 0 (por ejemplo) y luego se incrementa. Es crucial que los valores producidos por el contador nunca se repitan para evitar cualquier ambigüedad.

Las ventajas del contador radican en su simplicidad, lo que lo convierte en una buena opción para conjuntos de datos pequeños y no complejos. En términos de protección de datos, el contador proporciona seudónimos sin conexión con los identificadores iniciales (aunque el carácter secuencial del contador aún puede proporcionar información sobre el orden de los datos dentro de un conjunto de datos). Sin embargo, esta solución puede tener problemas de implementación y escalabilidad en casos de conjuntos de datos grandes y más sofisticados, ya que es necesario almacenar la tabla completa de mapeo de seudonimización.

Generador de Números Aleatorios (GNA)

El GNA es un mecanismo que produce valores en un conjunto que tienen una probabilidad igual de ser seleccionados de la población total de posibilidades y, por lo tanto, son impredecibles. Este enfoque es similar al contador, con la diferencia de que se asigna un número aleatorio al identificador.

Dos opciones están disponibles para crear este mapeo: un generador de números aleatorios verdaderos o un generador pseudoaleatorio criptográfico (Ver definiciones en [46]). Cabe señalar que, en ambos casos, sin el debido cuidado, pueden ocurrir colisiones. Una colisión es el caso en el que dos identificadores están asociados al mismo seudónimo. La probabilidad de que aparezca una colisión está relacionada con el conocido problema de cumpleaños [47] en el que se basa el ataque homónimo (Birthday Attack).

El GNA proporciona una sólida protección de datos (que, a diferencia del contador, utiliza un número aleatorio para crear cada seudónimo, por lo tanto, es difícil extraer información sobre el identificador inicial, a menos que se comprometa la tabla de mapeo). Las colisiones pueden ser un problema, como se mencionó anteriormente, al igual que la escalabilidad (la tabla completa de mapeo de seudonimización debe ser almacenada), dependiendo del escenario de implementación.

E-mail address	Pseudonym (Random number generator)	Pseudonym (counter generator)
alice@abc.eu	328	10
bob@wxyz.com	105	11
eve@abc.eu	209	12
john@qed.edu	83	13
alice@wxyz.com	512	14
mary@clm.eu	289	15

Figura 3-4 Ejemplo de contador y generador de números aleatorios [48].

Función Criptográfica de Hash

Una función criptográfica de hash toma cadenas de entrada de longitud arbitraria y las mapea a salidas de longitud fija. Estas funciones poseen las siguientes propiedades:

- **Unidireccional:** es computacionalmente inviable encontrar cualquier entrada que se mapee a una salida predefinida.
- **Libre de colisiones:** es computacionalmente inviable encontrar dos entradas distintas que se mapeen a la misma salida.

Una función criptográfica de hash se aplica directamente al identificador para obtener el seudónimo correspondiente: $Pseudo = H(Id)$. El valor hash producido por la función de hash tiene una longitud predefinida, la cual está determinada por la tecnología de hash empleada. En este proceso, el conjunto de caracteres del valor de entrada no tiene relevancia; sin importar cuáles sean, siempre se

generará un número fijo de caracteres en el resultado hash. Los caracteres que son válidos están definidos por las reglas establecidas en la función de hash.

Aunque una función de hash puede contribuir significativamente a la integridad de los datos, generalmente se considera débil como técnica de seudonimización ya que es propensa a ataques de fuerza bruta y de diccionario.

Código de Autenticación de Mensajes (CAM)

Esta técnica puede verse como una función de hash con clave. Es muy similar a la solución anterior excepto que se introduce una clave secreta para generar el seudónimo. Sin el conocimiento de esta clave, no es posible relacionar los identificadores y los seudónimos. HMAC [49] es, con mucho, el diseño de código de autenticación de mensajes más popular utilizado en los protocolos de Internet.

El CAM (MAC en inglés) generalmente se considera una técnica robusta de seudonimización desde el punto de vista de la protección de datos, ya que revertir el seudónimo se supone inviable, siempre y cuando la clave no haya sido comprometida. Diferentes variaciones del método pueden aplicarse según los requisitos de utilidad y escalabilidad de la entidad de seudonimización.

Cifrado

Se considera principalmente el **cifrado simétrico (determinista)** y, en particular, los cifradores por bloques como el AES y sus modos de operación [46]. El cifrador por bloques se utiliza para cifrar un identificador utilizando una clave secreta, que es tanto el secreto de seudonimización como el secreto de recuperación. El uso de cifradores por bloques para seudonimización requiere tratar con el tamaño del bloque. El tamaño de los identificadores puede ser más pequeño o más grande que el tamaño del bloque de entrada del cifrador por bloques. Si el tamaño de los identificadores es más pequeño, se debe considerar añadir un relleno. En el caso de que el tamaño de los identificadores sea mayor que el tamaño del bloque, existen dos opciones que se pueden utilizar para resolver este problema; los identificadores se pueden comprimir en algo más pequeño que el tamaño del bloque; si la compresión no es una opción disponible, se puede usar un modo de operación (como el modo de contador CTR). Sin embargo, esta última opción requiere manejar un parámetro adicional, el vector de inicialización.

El cifrado también puede ser una técnica robusta de seudonimización, con varias propiedades similares a las del CAM.

El **cifrado probabilístico** es otra alternativa que podría usarse especialmente en casos donde se necesite generar diferentes seudónimos para el mismo identificador.

Otro tipo de cifrado podría ser **el cifrado homomórfico.** Esta estrategia de "privacidad por defecto" resulta apropiada en situaciones en las que un responsable subcontrata una porción de una actividad a un encargado y busca asegurar de manera técnica que este último no podrá acceder a los datos [50].

Perturbación de datos

La perturbación de datos es una técnica que modifica los valores de un conjunto de datos original para que sean ligeramente diferentes. Se utiliza para proteger la privacidad de los datos sensibles cuando se utilizan identificadores indirectos, como números o fechas. Estos identificadores pueden ser potencialmente identificables cuando se combinan con otras fuentes de datos, pero los cambios leves en el valor no afectan significativamente al atributo [51].

La perturbación de datos no debe utilizarse cuando la precisión de los datos sea crucial. Por ejemplo, no se debe utilizar para datos financieros o médicos.

El método exacto de perturbación de datos utilizado depende del tipo de atributo que se desea proteger. Los métodos comunes incluyen:

- Redondeo: Se redondea el valor del atributo a una base determinada, como 10 o 100. Por ejemplo, un número de teléfono de 10 dígitos podría redondearse a un número de 5 dígitos.
- Adición de ruido aleatorio: Se agrega ruido aleatorio al valor del atributo. Por ejemplo, un número de edad podría aumentarse o disminuirse aleatoriamente en unos pocos años.

El grado de perturbación debe ser proporcional al rango de valores del atributo. Si la base es demasiado pequeña, el efecto de anonimización será más débil; por otro lado, si la base es demasiado grande, los valores finales serán demasiado diferentes del original y la utilidad del conjunto de datos probablemente se reducirá. Cuando el cálculo se realiza sobre valores de atributos que se han perturbado antes, el valor resultante puede experimentar perturbaciones en una medida aún mayor.

3.2.4.2 Políticas de Seudonimización

Mientras que la elección de una técnica de seudonimización es esencial, la política (o modo) de implementación de la seudonimización es igualmente importante para su aplicación práctica. En esta parte, se considera el problema más general de la seudonimización de una base de datos o cualquier documento que contenga k identificadores. Consideremos un identificador Id que aparece varias veces en dos conjuntos de datos A y B. Después de la seudonimización, el identificador Id se sustituye según una de las siguientes políticas: seudonimización determinista, seudonimización con aleatorización de documentos y seudonimización completamente aleatorizada.

<u>Seudonimización determinista</u>

En todas las bases de datos y cada vez que aparece, Id siempre se reemplaza por el mismo seudónimo $pseudo$. Es consistente dentro de una base de datos y entre diferentes bases de datos. El primer paso para implementar esta política es extraer la lista de identificadores únicos contenidos en la base de datos. Luego, esta lista se asigna a los seudónimos y finalmente los identificadores se sustituyen por los seudónimos en la base de datos.

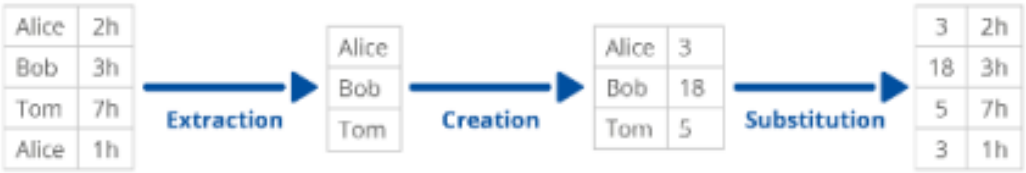


Figura 3-5 Ejemplo de Seudoanonimización determinista [48]

<u>Seudonimización con aleatorización de documentos</u>

Cada vez que Id aparece en una base de datos, se sustituye por un seudónimo diferente ($pseudo1$, $pseudo2$, etc). Sin embargo, Id siempre se asigna a la misma colección de ($pseudo1$, $pseudo2$) en los conjuntos de datos A y B. La seudonimización es consistente solo entre diferentes bases de datos en este caso. La tabla de mapeo se crea esta vez utilizando todos los identificadores contenidos en la base de datos. Cada ocurrencia de un identificador dado (por ejemplo, Alice) se trata de forma independiente.

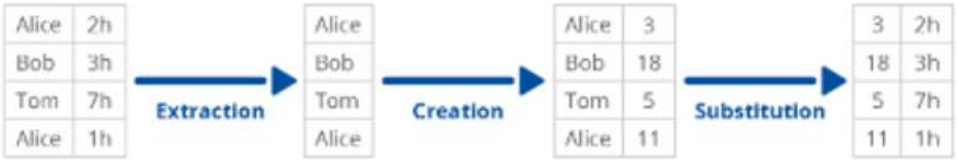


Figura 3-6 Ejemplo de Seudoanonimización con aleatoriedad de documentos [48]

Seudonimización completamente aleatorizada

Finalmente, para cualquier ocurrencia de Id dentro de una base de datos A o B, Id se sustituye por un seudónimo diferente ($pseudo1$, $pseudo2$). Este caso es la seudonimización completamente aleatorizada. Esta política se puede ver como una extensión adicional de la "seudonimización con aleatorización de documentos". De hecho, las dos políticas tienen el mismo comportamiento cuando se aplican en un solo documento. Sin embargo, si el mismo documento se seudonimiza dos veces con seudonimización completamente aleatorizada, se obtienen dos salidas diferentes. Con la seudonimización con aleatorización de documentos, se habría obtenido la misma salida dos veces. En otras palabras, en la seudonimización con aleatorización de documentos, la aleatoriedad es selectiva (por ejemplo, solo para Alice), mientras que en la seudonimización completamente aleatorizada, la aleatoriedad es global (se aplica a cualquier registro).

3.2.4.3 Técnicas de Seudonimización Avanzadas

Además de las técnicas de seudonimización mencionadas anteriormente, existen una gran variedad de alternativas más avanzadas, adecuadas a contextos diversos.

Adicionalmente al simple hash de datos, se pueden utilizar estructuras más avanzadas como los árboles de Merkle [52], que utilizan hashes de conjuntos de hashes, por ejemplo, h3=hash(h1,h2), para lograr seudónimos estructurados que solo pueden ser descubiertos parcialmente en lugar de completamente. Del mismo modo, las cadenas de hash [53] se basan en el hash repetido de los valores de hash de los valores de hash, por ejemplo, h4=h3(h2(h1(x))), para producir un valor que requiere múltiples inversiones de hash para volver a identificar los datos originales de un seudónimo dado. Un ejemplo de esta técnica de hash, una cadena de seudonimización, implica varias entidades de seudonimización que toman sucesivamente los seudónimos creados por la entidad de seudonimización anterior como entrada para crear nuevos seudónimos (por ejemplo, aplicando otra capa de hash). Tal cadena será válida incluso si un adversario logra descubrir todos los seudonimizaciones aplicadas en la cadena total, lo que la convierte en una técnica de seudonimización muy sólida. Es una práctica común, por ejemplo, en ensayos clínicos.

Si el dominio de entrada abarca múltiples dimensiones, los filtros de Bloom [54], además de ser utilizados como técnica de anonimización, se pueden utilizar para realizar eficientemente una seudonimización computacionalmente factible sobre todas las posibles combinaciones de valores de entrada en los diferentes dominios, a pesar del problema de explosión de estados.

Los seudónimos de transacciones vinculables y/o la vinculación controlada de seudónimos con la opción de reidentificación gradual también pueden constituir otra opción a considerar [55].

Finalmente, todas las técnicas que pueden utilizarse de manera efectiva para aumentar la anonimización también pueden ser útiles para la seudonimización, como las técnicas comunes para k-anonimato o la privacidad diferencial. Las pruebas de conocimiento nulo [56] y el área más amplia de credenciales basadas en atributos también pueden proporcionar soluciones a tener en cuenta [57].

La elección de una técnica de seudonimización depende de diferentes parámetros, como el nivel de protección de los datos y la utilidad del conjunto de datos seudonimizado que la entidad de seudonimización desea lograr. En términos de protección, los generadores de números aleatorios (GNA), los códigos de autenticación de mensajes (CAM) y el cifrado son técnicas más sólidas, ya que por diseño impiden la búsqueda exhaustiva, la búsqueda en diccionarios y las conjeturas. No obstante, los requisitos de utilidad pueden llevar a la entidad de seudonimización a considerar una combinación de enfoques diferentes o variaciones de un enfoque seleccionado.

De manera similar, la seudonimización completamente aleatorizada ofrece el nivel de protección más alto, pero impide cualquier comparación entre bases de datos. Las funciones de seudonimización con aleatorización de documentos y las funciones determinísticas proporcionan utilidad, pero permiten la vinculación entre registros.

3.3 Desafíos y limitaciones.

Como hemos definido anteriormente, la anonimización de datos es un proceso utilizado para eliminar o modificar información identificable en conjuntos de datos, con el objetivo de proteger la privacidad, haciendo que sea más difícil identificar a los individuos a partir de los datos y cumplir con las regulaciones de privacidad.

La anonimización puede ser difícil de implementar de manera efectiva. Es importante elegir la técnica de anonimización adecuada para el tipo de datos que se están tratando y para el nivel de privacidad que se desea lograr

El proceso de anonimización, lejos de ser perfecto, presenta una serie de desafíos y limitaciones importantes.

Desafíos:

Entre los principales riesgos y retos asociados a la anonimización existentes, podemos nombrar los siguientes [58] :

- El avance de la tecnología, que hace especialmente complejo poder garantizar la anonimización absoluta.
- Reidentificación: Existe el riesgo de reidentificación, donde datos anonimizados pueden ser combinados con otras fuentes para identificar a individuos.
- Pérdida de utilidad: Al anonimizar datos, es posible que se elimine información valiosa y útil, lo que puede afectar la calidad y la utilidad del conjunto de datos para su propósito original.
- Mantenimiento de la utilidad: Equilibrar la protección de la privacidad con la necesidad de mantener la utilidad del conjunto de datos puede ser un desafío, ya que a veces la anonimización puede ser demasiado agresiva.
- Cambios en el contexto: Los cambios en el contexto o la adición de nuevos datos pueden afectar la eficacia de la anonimización original.
- Requerimientos legales y éticos: Las leyes y regulaciones de privacidad pueden ser complejas y variar según la jurisdicción, lo que dificulta garantizar el cumplimiento total y entender completamente las implicaciones legales y éticas.
- Problemas de obtención del consentimiento.
- Ejercicio de los derechos de información y acceso, rectificación, cancelación u oposición.
- La cuestión de la correcta anonimización de los datos antes de analizarlos.
- Desvío de la finalidad para la que fueron recogidos los datos y otras posibles vulneraciones del principio de calidad de los mismos (datos inexactos o excesivos).

Limitaciones:

- Complejidad de datos: Algunos tipos de datos, como secuencias temporales o datos geoespaciales, pueden ser más difíciles de anonimizar sin perder su utilidad.
- Granularidad variable: La granularidad variable de los datos puede afectar la eficacia de la anonimización, ya que algunos datos pueden ser más difíciles de desidentificar que otros.
- Escalabilidad: Anonimizar grandes conjuntos de datos puede ser un proceso costoso en términos de recursos y tiempo.
- Conocimiento previo: En algunos casos, si un atacante tiene conocimiento previo sobre los individuos en el conjunto de datos, puede ser más fácil realizar ataques de reidentificación.

- Necesidad de actualización constante: Dado que las técnicas de anonimización pueden volverse obsoletas con el tiempo debido a avances en la tecnología, es necesario actualizar constantemente las prácticas y técnicas utilizadas.

Es esencial abordar estos desafíos y limitaciones de manera cuidadosa y considerada para garantizar la eficacia de la anonimización y la protección de la privacidad de los individuos.

4 OCULTACIÓN DE LA HUELLA DIGITAL

4.1 Importancia de la ocultación de la huella digital.

La información digital se ha convertido en parte de la sociedad y con la expansión de Internet, la infraestructura de red se ha vuelto una parte indispensable en la sociedad, la industria, el ocio, el comercio y otras áreas de la actividad humana. Sin embargo, por diversas razones, las redes actuales son vulnerables a numerosos riesgos, como la fuga de información, la violación de la privacidad y la corrupción de datos [59].

Debido a las evidencias de la vigilancia electrónica masiva y a la recolección ilícita de datos personales, muchos usuarios de Internet recurren al uso de herramientas de comunicación anónimas y privadas [60].

Como consecuencia de las situaciones descritas anteriormente, los investigadores han diseñado sistemas de anonimato que construyen una capa de red que corre sobre Internet. Usando sistemas de anonimato, un usuario puede comunicarse con otro sin revelar su identidad, dirección IP o ubicación [61].

4.2 Métodos y técnicas de ocultación.

Para evitar la recopilación de información de un individuo por parte de un tercero a través de Internet, podemos por un lado utilizar las conocidas como redes de comunicación anónimas y por otro lado tratar de minimizar la huella digital poniendo en práctica una serie de medidas o acciones catalogadas como buenas prácticas.

4.3 Redes de comunicación anónimas

Las redes de comunicación anónimas (ACN – Anonymous Communication Network) son sistemas diseñados para permitir la comunicación en línea de manera completamente anónima y privada, ocultando la identidad del remitente, del destinatario y, a menudo, el contenido del mensaje. Estas redes son especialmente importantes para personas que desean proteger su privacidad, como periodistas, activistas políticos y ciudadanos en países con regímenes opresivos.

La comunicación anónima protege las identidades del remitente y del destinatario de terceros y mantiene oculta la identidad del usuario frente a usuarios remotos [62].

Aunque existen muchos algoritmos de encriptación, ocultar la identidad del remitente solo puede lograrse a través de una red anónima. En el diseño de las redes ACN, se definen las necesidades específicas de comunicación que se deben abordar para, de esta manera, concretar las características del sistema.

Existen diferentes clasificaciones de redes anónimas según las necesidades requeridas:

- Nivel de latencia:
  - Las redes anónimas de alta latencia son redes que introducen un retraso significativo en la transmisión de los datos, normalmente por razones de seguridad o privacidad. Por ejemplo, una red de alta latencia puede enviar los datos a través de varios nodos intermedios que los mezclan y los reenvían de forma aleatoria, lo que dificulta el seguimiento y el análisis de la red.
  - Las redes anónimas de baja latencia [63] son aquellas que minimizan el retraso en la transmisión de los datos, normalmente por razones de rendimiento o funcionalidad. Por ejemplo, una red de baja latencia puede utilizar técnicas de cifrado y enrutamiento que reducen el número de saltos y la sobrecarga de los datos, lo que mejora la velocidad y la eficiencia de la red.
- Nivel de anonimato:
  - Se define nivel de anonimato como la probabilidad que un intruso pueda identificar a un componente de la comunicación como el emisor original de un mensaje [64].
  - La necesidad cuantificar en nivel de anonimato no es una cuestión trivial y se ha considerado como un gran reto en el que se ha desarrollado una importante labor de investigación desde los inicios [65] de las redes ACN. Debido a la importancia de la correcta medida o cuantificación del anonimato, se han desarrollado multitud de métricas y metodologías intentando dar una solución al problema planteado.

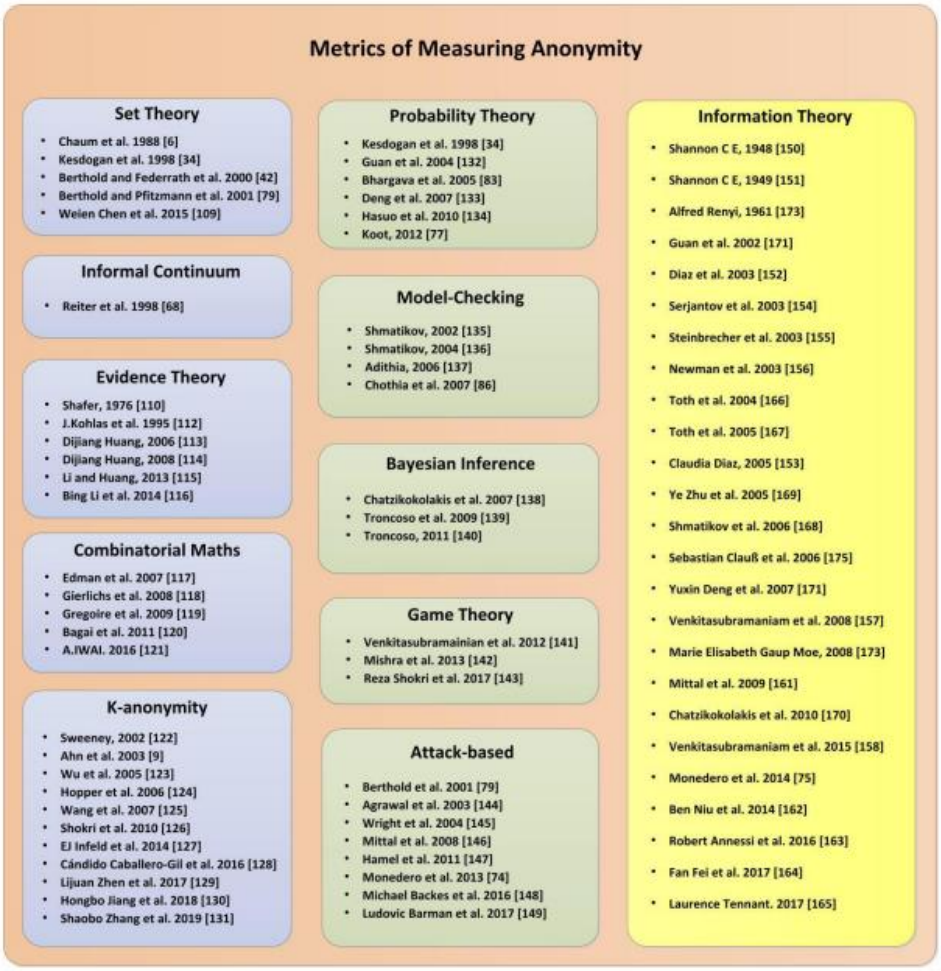


Figura 4-1 Resumen de diferentes métricas para medir al anonimato [41].

- Tipo de infraestructura (centralizada o distribuida)

En resumen, podemos decir que existen multitud de redes de comunicación anónima con diferentes características y funcionalidades en función de los mecanismos de anonimato que utilicen.

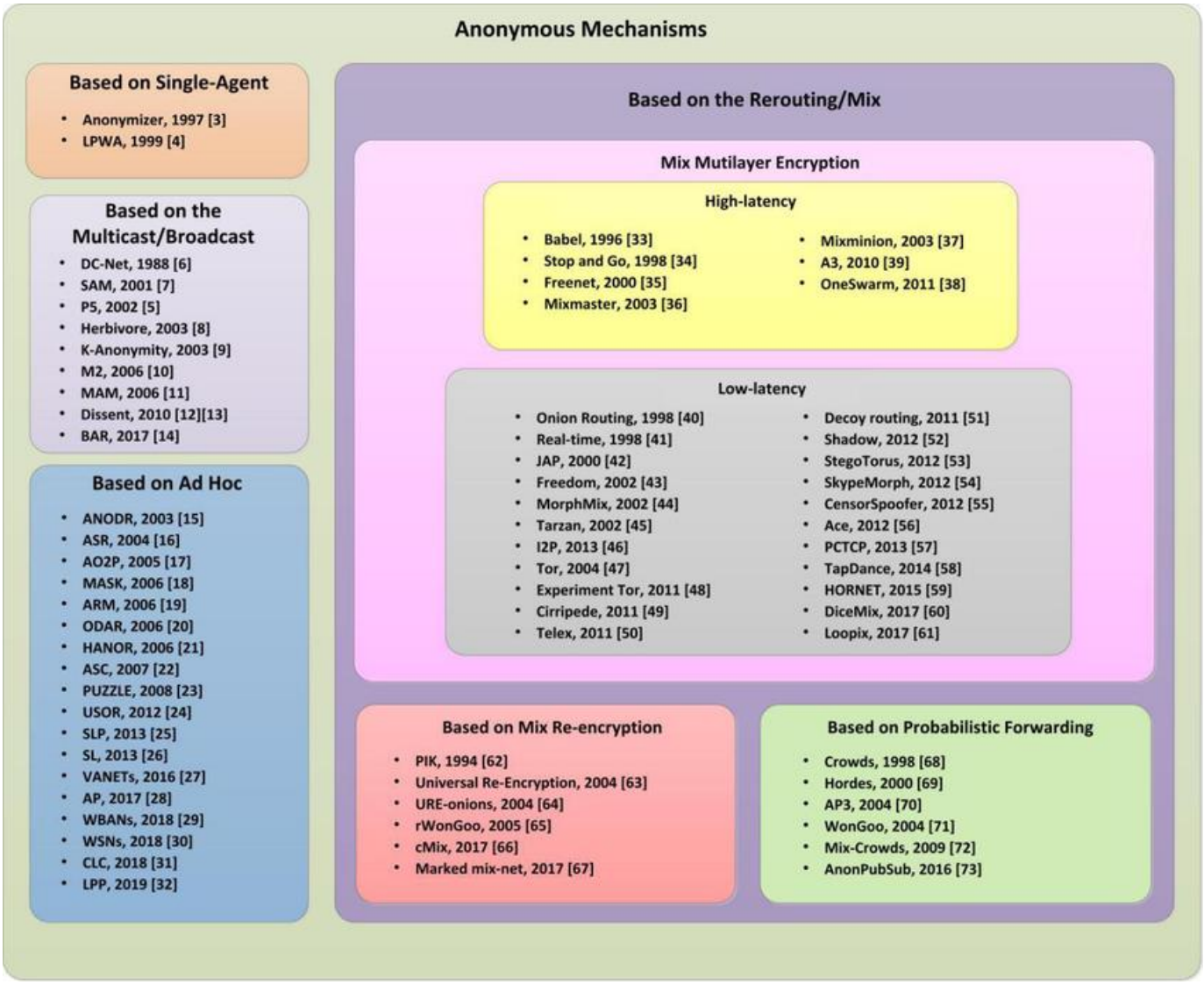


Figura 4-2 Muestra de redes anónimas basadas en diferentes mecanismos de anonimato [66].

4.3.1 Las primeras redes anónimas. Mix Network

<u>Mix Network</u>

En 1981, David Chaum publicó una propuesta de red de comunicación anónima [67]. Su propuesta, llamada Mix Network, permite a un grupo de remitentes enviar un mensaje encriptado con su destinatario a un servidor. Una vez que el servidor tiene una cantidad suficiente de mensajes (lote), los reordena y ofusca para que solo este servidor sepa qué mensajes proviene de qué remitente. El lote se reenvía a otro servidor que realiza el mismo proceso. Finalmente, los mensajes llegan al servidor final, donde se descifran por completo y se envían al destinatario. También, Chaum propone un mecanismo para permitir el retorno de los mensajes. Las Mix Networks son la base de los conocidos servidores "remailers"[1] (pseudoanónimos, cypherpunk, mixmaster, mixminion, etc) y además son el

[1] Un "remailer" o reenviador de correo anónimo es un servidor que recibe mensajes con instrucciones incrustadas sobre dónde enviar dichoa mensaje a continuación y reenvíarlos sin revelar su origen.

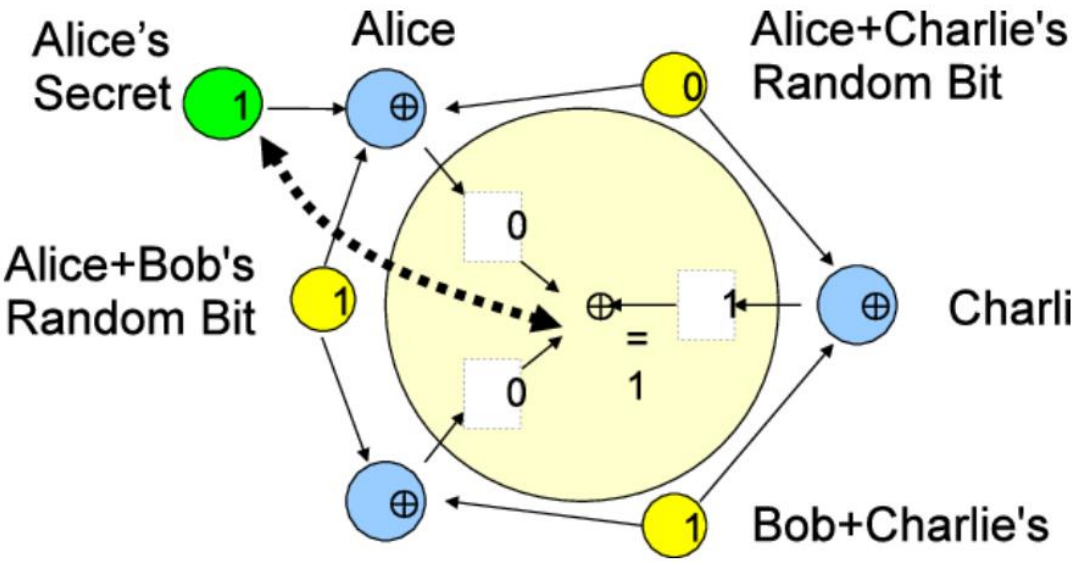
antecesor conceptual de las herramientas modernas de navegación web anónima como Tor (basadas en el enrutamiento cebolla). Chaum ha abogado porque cada enrutador se convierta, en efecto, en un nodo Tor.

DC-Nets

En 1988, Chaum da un paso más e introduce un tipo diferente de sistema de comunicación anónima conocido por DC-Net, el cual es una solución al problema conocido como la cena de los criptógrafos [68].

La solución al problema demuestra que es posible crear un sistema de comunicación en el que los participantes puedan comunicarse entre sí sin que se pueda rastrear el origen de los mensajes. Este sistema también es seguro contra ataques computacionales [68].

Las DC-nets son un enfoque criptográfico que permite a un grupo de usuarios compartir información de manera anónima. Cada usuario agrega ruido a sus datos personales antes de compartirlos, lo que hace que sea difícil rastrear la información a un individuo específico. Solo el propietario original de los datos puede eliminar el ruido y recuperar la información original.



The Dining Cryptographers approach to anonymous communication. Alice reveals a 1-bit secret to the group, but neither Bob nor Charlie learn which of the other two members sent this message.

Figura 4-3 DC-Net [69]

El funcionamiento de las DC-nets, en su forma más simple, ilustrada en la Figura 4-3, asumimos que un miembro del grupo desea transmitir de manera anónima un mensaje de 1 bit. Para lograrlo, cada par de miembros lanza una moneda, acordando en secreto el resultado aleatorio de ese lanzamiento de moneda. Un grupo de N miembros lanza así un total de N(N−1)/2 monedas, de las cuales cada miembro observa el resultado de N−1 monedas. Luego, cada miembro realiza la operación XOR en los valores de las N−1 monedas que observa, además el miembro que desea transmitir el mensaje de 1 bit realiza la operación XOR con el valor de ese mensaje, para producir el texto cifrado DC-nets de ese miembro. Cada miembro del grupo luego transmite su texto cifrado de 1 bit a los demás miembros. Finalmente, cada miembro recopila y realiza la operación XOR en todos los textos cifrados de los N miembros juntos. Dado que el valor de cada moneda compartida se le aplica la operación XOR exactamente en los textos cifrados de dos miembros, todas las monedas se cancelan, dejando solo el mensaje anónimo, revelando de manera demostrable ninguna información sobre qué miembro del grupo envió el mensaje.

Por extensión, en las DC Net se envía por cada estación participante en la comunicación un mensaje real o ficticio en un momento determinado y la combinación de estos mensajes es recibida por

todas las estaciones. Cada estación genera claves aleatorias que comparte con las demás estaciones de la red. Así, cada estación tiene n-1 claves. Si una estación quiere enviar un mensaje, hace una combinación local, es decir, suma el mensaje, las claves que generó y el inverso de las claves que recibió. Si una estación no quiere enviar un mensaje, debe enviar un mensaje vacío, combinado con todas las claves que conoce. Todas las salidas se combinan globalmente y se distribuyen a cada una de las estaciones de la red. Como cada clave y su inverso se sumaron solo una vez, las claves se cancelan entre sí en la combinación global. Además, el resultado de la combinación global es la suma de todos los mensajes enviados. Si ninguna estación quiere enviar un mensaje, la suma será un mensaje vacío. Si solo una estación envía un mensaje, la suma será igual a ese mensaje. Si dos o más estaciones envían sus mensajes al mismo tiempo, sus resultados combinados podrían provocar colisiones, lo que es un problema en los sistemas de distribución en canales con accesos múltiples. Al final, cada participante en el sistema tiene el mensaje original [70].

cMix - PrivaTegrity

En 2017, Chaum publica una descripción de una nueva variedad de Mix Network. Son las redes cMix, diseñadas para ser mucho más eficientes que el esquema de encriptación por capas propuesto anteriormente.

La implementación de la red cMix, se ha dado a conocer en la nueva red de comunicación llamada PrivaTegrity [71], que tiene por objetivo solventar las debilidades de la red TOR. Está basada en una arquitectura de nueve servidores privados, con un décimo que actúa como gestor y que no tiene acceso a ninguna clave secreta.

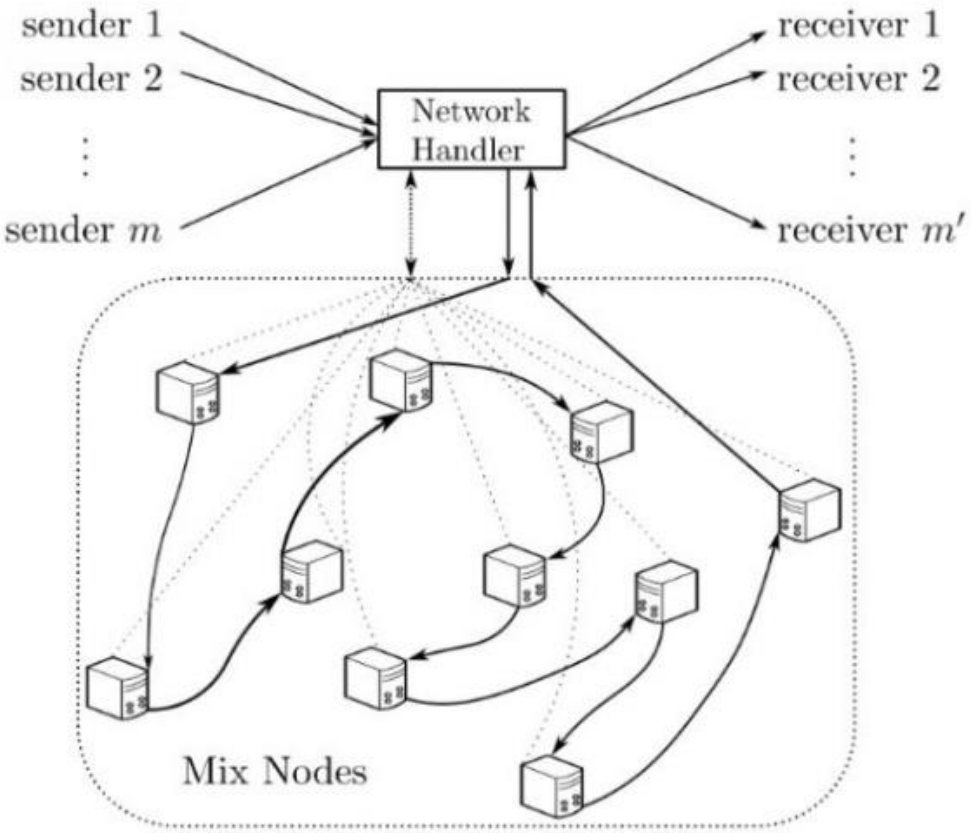


Figura 4-4 PrivaTegrity [72].

El funcionamiento de esta red es el siguiente: Durante la instalación y configuración de cMix, un teléfono inteligente (smartphone) se comunica con los nueve servidores de PrivaTegrity estableciendo una serie de claves únicas que comparte con cada servidor. Cuando el teléfono envía un mensaje, cifra los datos del mensaje aplicando una serie de claves únicas creadas. Después, el mensaje se transmite

entre los nueve servidores de uno en uno, cada uno de los cuales particiona el mensaje con su clave secreta y aplica a los datos resultantes un número aleatorio. Entonces, se realiza una segunda transmisión entre los nueve servidores, el mensaje se coloca en una pila con otros mensajes, y cada servidor desordena el orden de la pila utilizando un patrón aleatorio que solo ese servidor conoce, y luego aplica a la pila de los mensajes otro número aleatorio. Finalmente, el proceso se revierte, y a medida que el mensaje pasa por los servidores por última vez, todos esos números aleatorios se anulan y se reemplazan por las claves únicas del destinatario del mensaje, quien puede entonces descifrarlo y leerlo [73]

4.3.2 El proyecto TOR

El Proyecto Tor ha pasado por varias generaciones de desarrollo para mejorar la privacidad, la seguridad y anonimato de los usuarios. Se pueden distinguir tres generaciones del enrutamiento routers cebolla (OR – Onion Routers) [74]:

Generación 0 (1995-2004):

El concepto de enrutamiento de cebolla (onion routing) fue desarrollado por el Laboratorio de Investigación Naval de los Estados Unidos a mediados de la década de 1990. Los enrutadores de cebolla de primera generación se utilizaron con fines de investigación y no estaban disponibles públicamente. Este diseño tenía algunas limitaciones, como la falta de secreto directo perfecto, el control de congestión, los servidores de directorio y la verificación de integridad.

Primera Generación (2004-2009):

El Proyecto Tor fue fundado en 2004 y se lanzó la primera versión pública de Tor. Esto marcó el comienzo de la primera generación de enrutadores Tor. Durante este período, Tor se hizo ampliamente utilizado y su base de usuarios creció rápidamente. Esta generación corresponde a todas las mejoras del diseño que se hicieron después de la generación 0 y antes de Tor. Estas mejoras incluían el uso de cifrado asimétrico, la introducción de políticas de salida configurables y el diseño de servicios ocultos mediante puntos de encuentro.

Segunda Generación (2009-Presente):

La segunda generación de enrutadores Tor, corresponde a TOR propiamente dicho, es el acrónimo de The Onion Router. Tor es un software libre que se lanzó en 2002 y que se basa en el diseño de la primera generación, pero con algunas modificaciones y optimizaciones realizadas a través de desarrollos continuos que mejoran la seguridad, la escalabilidad y la facilidad de uso. Un avance significativo en esta generación fue la introducción de "pluggable transports", que permiten a Tor eludir la censura de la red al disfrazar el tráfico de Tor como otros tipos de tráfico de Internet. También se introdujo durante este período el Navegador Tor, una versión modificada de Mozilla Firefox, que facilita a los usuarios acceder a la red Tor de forma segura.

4.3.3 TOR

La red TOR es una red de anonimato y privacidad en línea que permite a sus usuarios acceder a sitios web que no están disponibles en la red convencional. TOR es el acrónimo de The Onion Router, o el router cebolla, porque emplea un sistema de encriptación de varias capas que dificulta enormemente el seguimiento de la identidad y la localización de los usuarios [75].

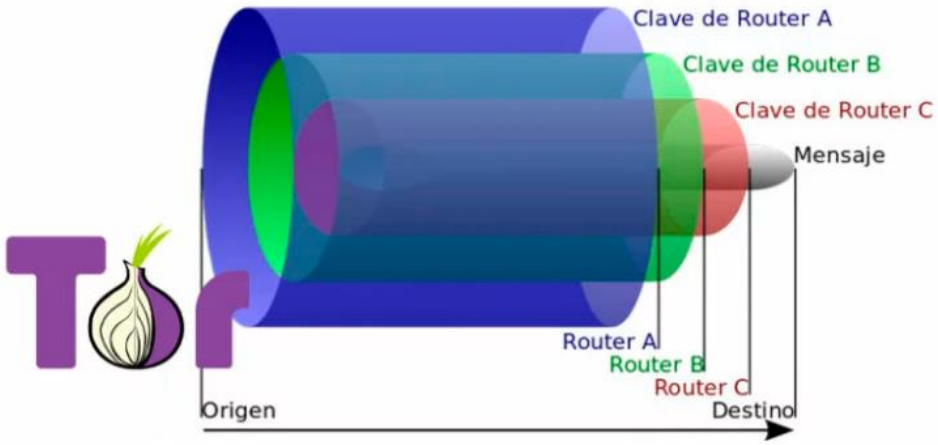


Figura 4-5 Encapsulamiento red TOR [76]

La red Tor se basa en un diseño de enrutamiento por cebolla [77], donde el tráfico se reenvía a través de varios routers y se cifra de forma múltiple, en cada router se elimina una capa del cifrado. El camino a través de la red, conocido como túnel, se construye de forma telescópica, de modo que cada router solo conoce el router anterior y el siguiente en el camino. En particular, el primer router (nodo de entrada) conoce el origen del túnel, pero no su destino, y el último router (nodo de salida) conoce el destino, pero no el origen [78].

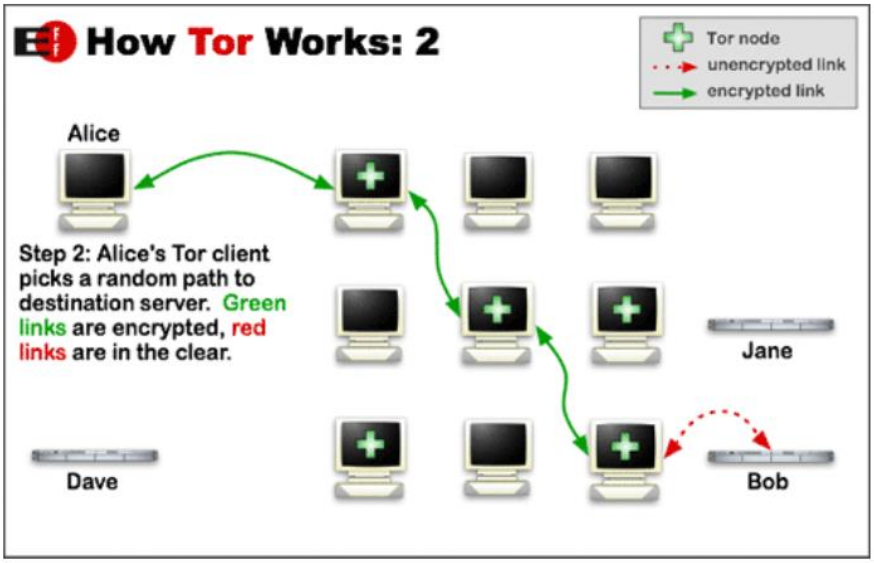


Figura 4-6 Red TOR [79]

4.3.4 Redes peer to peer (P2P)

Las redes peer-to-peer (P2P), o redes de pares, han surgido como un modelo de negocio y una arquitectura de sistemas viable para aplicaciones distribuidas en Internet. Aunque sus raíces tecnológicas se remontan varias décadas y se fundamentan en el diseño de sistemas de información distribuidos, las aplicaciones actuales demuestran que es una forma eficaz de construir aplicaciones que conectan a millones de usuarios en todo el mundo sin necesidad de desplegar servidores dedicados. En su lugar, combinan los recursos de los ordenadores de cada usuario, permitiendo que estos sistemas se autoorganicen y adapten automáticamente a los cambios en el número de usuarios intercambiados en cada momento, a la vez que proporcionan servicios para el intercambio de contenidos y las comunicaciones personales [80].

Los sistemas P2P han sido utilizados en diferentes dominios de aplicaciones [81], sin embargo, su arquitectura adquirió fama y se generalizó su conocimiento por el sistema de archivos Napster, lanzado inicialmente en 1999 [82]. Napster marcó el inicio de las redes peer-to-peer, tal y como las conocemos actualmente, donde los usuarios que participan crean una red virtual, totalmente autónoma de la red física, sin tener que someterse a ninguna autoridad o limitación administrativa [83].

El sistema P2P permitió a una gran cantidad de usuarios conectarse a través de Internet "directamente, formando grupos y colaborando para convertirse en motores de búsqueda, supercomputadoras virtuales y sistemas de archivos creados por los usuarios"[84].

Un sistema de comunicación P2P anónimo es una aplicación distribuida entre pares en la que los nodos, que se utilizan para compartir recursos, o los participantes son anónimos o seudónimos.[85]. El anonimato de los participantes suele lograrse mediante redes de superposición de enrutamiento especiales que ocultan la ubicación física de cada nodo a los demás participantes [86].

La idea de Tim Berners-Lee para la World Wide Web se inspiraba en una red P2P, ya que suponía que cada usuario de la web sería un creador y participante activo, generando y conectando contenido para formar una "web" de enlaces interrelacionados. La primera Internet era más libre que la actual, donde dos ordenadores conectados a Internet podían intercambiar paquetes entre sí sin firewalls ni otras medidas de seguridad [84].

La P2P tiene una arquitectura de aplicación distribuida que divide las tareas o las cargas de trabajo entre la red de ordenadores. Los ordenadores de la red funcionan sin clientes ni servidores fijos y todos tienen los mismos privilegios y son equipotentes en la red. Los ordenadores de la red pueden trabajar de forma simultánea como clientes y servidores en relación con los otros ordenadores de la red. De esta manera, estas redes permiten el intercambio de información, en cualquier formato, entre los diferentes nodos interconectados de la misma [87].

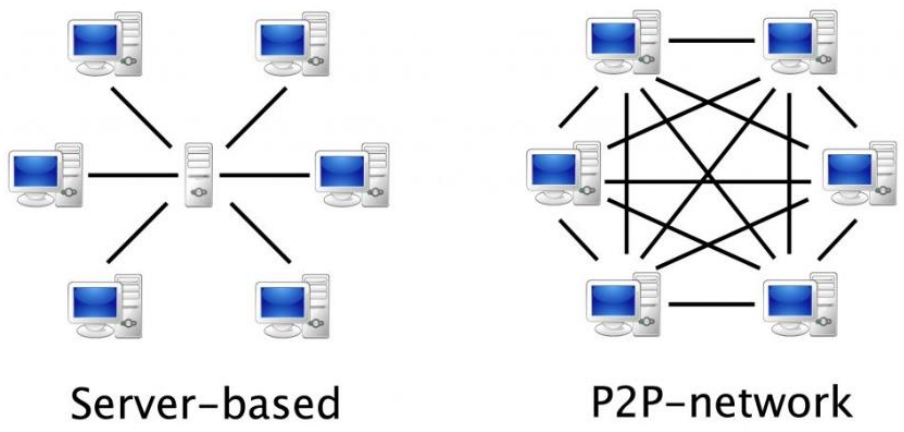


Figura 4-7 Arquitectura Cliente-Servidor vs P2P [80].

Las redes P2P se caracterizan por ser [88]:

- Escalables: Pueden soportar un número muy grande de usuarios sin necesidad de invertir en más servidores.
- Robustas: No dependen de un único servidor central, por lo que la caída de un nodo no implica la caída total del sistema.
- Descentralizadas: No tienen un servidor central, sino que todos los nodos son iguales y pueden actuar como clientes o servidores.

A nivel de seguridad, las redes peer-to-peer (P2P) no tienen ningún servidor especial para autenticar usuarios. Cada ordenador administra su propia seguridad, por lo que es posible que se deba crear una cuenta de usuario independiente para todo ordenador al que un usuario necesite acceder. Los usuarios suelen almacenar los archivos en sus propios ordenadores y son responsables de garantizar que se realicen copias de seguridad adecuadas de esos archivos. En una red P2P, cada ordenador suele ejecutar software de cliente y de servidor, y puede utilizarse para poner recursos a disposición de otros usuarios o para acceder a recursos compartidos en la red [89].

Algunas de las desventajas de las redes P2P son una seguridad débil, la falta de un almacenaje de archivos centralizado y la gestión de las copias de seguridad.

Las redes P2P, suelen implementar alguna forma de red virtual superpuesta sobre la topología de la red física, donde los nodos de la superposición forman un subconjunto de los nodos de la red física. Los datos se siguen intercambiando directamente sobre la red TCP/IP subyacente, pero en la capa de aplicación los pares pueden comunicarse entre sí directamente, a través de los enlaces lógicos de superposición (cada uno de los cuales corresponde a una ruta a través de la red física subyacente) [90]. Las superposiciones se utilizan para la indexación y el descubrimiento de pares y hacen que el sistema P2P sea independiente de la disposición de la red física.

4.3.4.1 Clasificación redes P2P

Durante su evolución, las redes P2P han sido clasificadas de múltiples maneras. Como muestra, se indican diferentes clasificaciones que aparecen en la literatura relacionada.

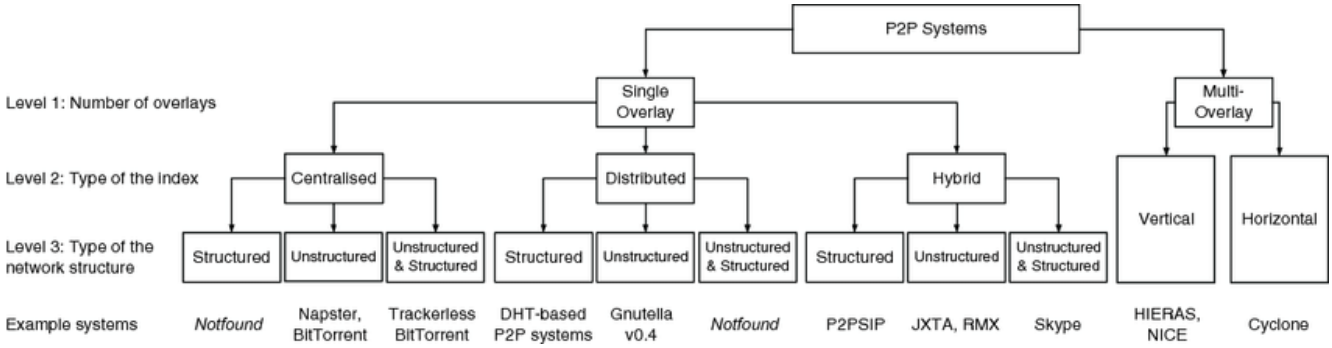


Figura 4-8 Clasificación redes P2P [91]

Otro tipo de clasificación podría ser la siguiente:

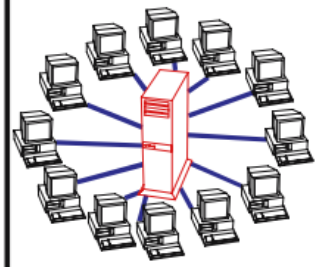
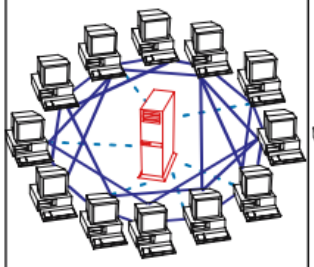
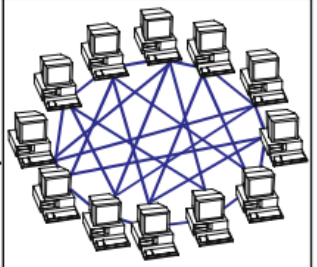
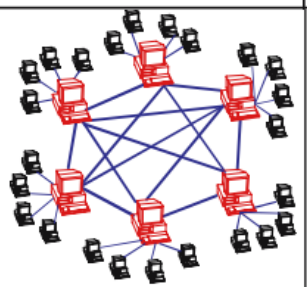
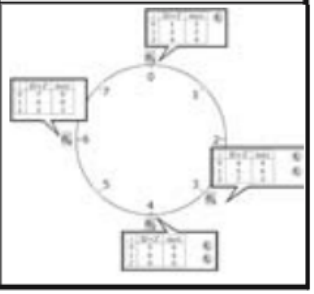
Client-Server	**Peer-to-Peer**			
	1. Resources are shared between the peers 2. Resources can be accessed directly from other peers 3. Peer is provider and requestor (Servent concept)			
	Unstructured P2P			**Structured P2P**
	1st Generation		**2nd Generation**	
1. Server is the central entity and only provider of service and content. → Network managed by the Server 2. Server as the higher performance system. 3. Clients as the lower performance system Example: WWW	*Centralized P2P*	*Pure P2P*	*Hybrid P2P*	*DHT-Based*
	1. All features of Peer-to-Peer included 2. Central entity is necessary to provide the service 3. Central entity is some kind of index/group database Example: Napster	1. All features of Peer-to-Peer included 2. Any terminal entity can be removed without loss of functionality 3. → No central entities Examples: Gnutella 0.4, Freenet	1. All features of Peer-to-Peer included 2. Any terminal entity can be removed without loss of functionality 3. → dynamic central entities Example: Gnutella 0.6, JXTA	1. All features of Peer-to-Peer included 2. Any terminal entity can be removed without loss of functionality 3. → No central entities 4. Connections in the overlay are "fixed" Examples: Chord, CAN
				

Figura 4-9 Clasificación redes P2P [92]

Como último ejemplo mostramos la siguiente clasificación:

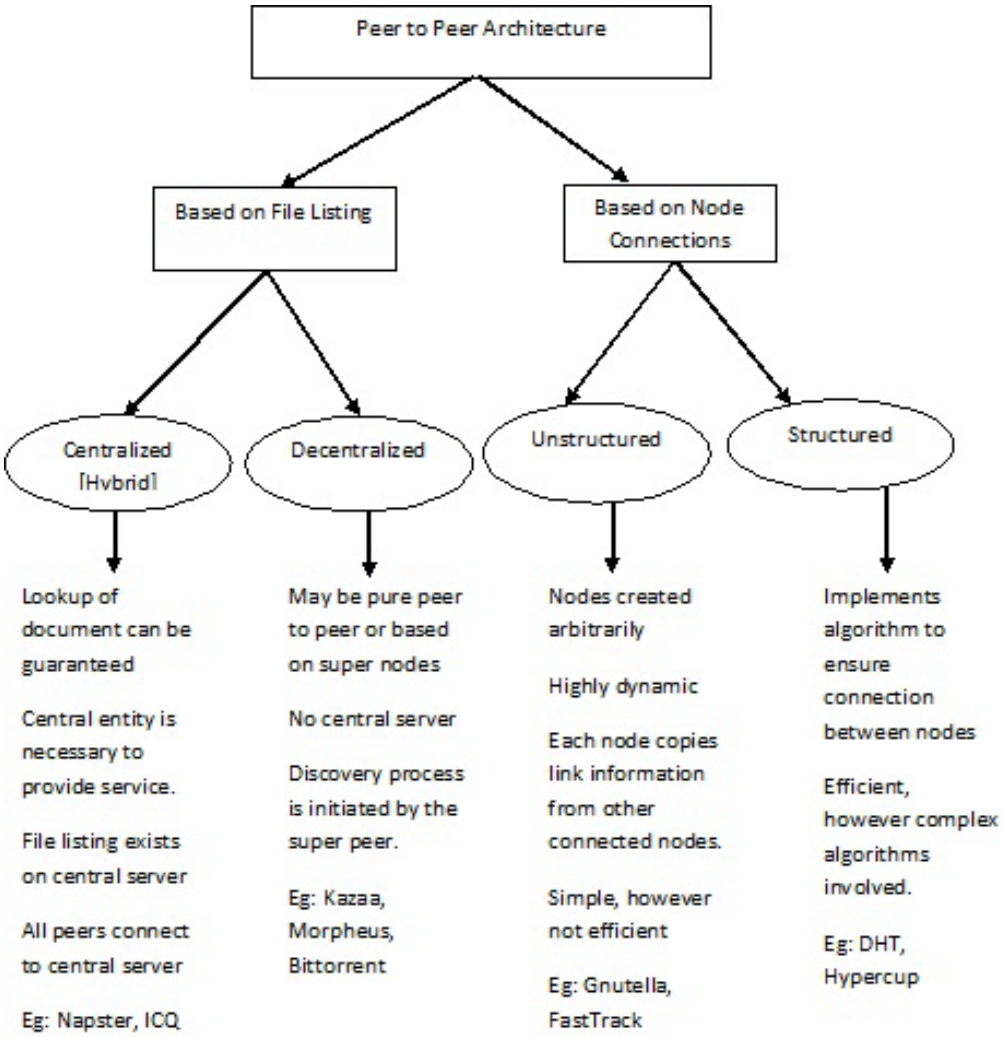


Figura 4-10 Clasificación redes P2P [93]

Podemos observar que en todas las clasificaciones de los sistemas P2P hay puntos comunes con lo que pasamos a describir con más detalle las redes estructuradas y las no estructuradas.

Redes Estructuradas o no estructuradas

Según cómo se conecten los nodos entre sí dentro de la red de superposición, y cómo se indexen y localicen los recursos, podemos clasificar las redes como no estructuradas o estructuradas (o como un híbrido entre ambas).

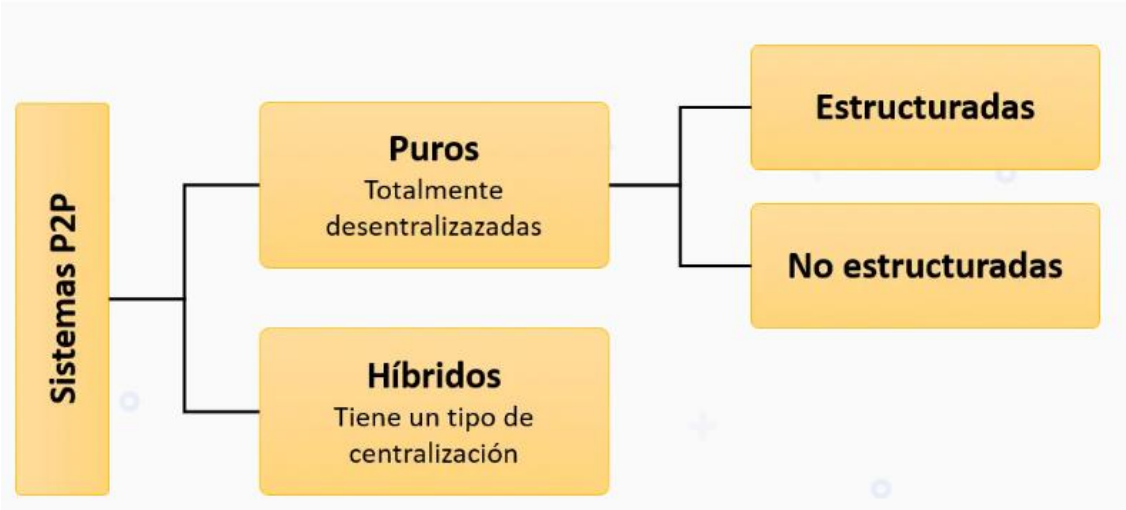


Figura 4-11 Clasificación redes P2P [94]

Las redes peer-to-peer **no estructuradas** son las más extendidas y se utilizan para una amplia gama de aplicaciones, incluyendo el intercambio de archivos, los juegos en línea y la distribución de contenido.

En las redes peer-to-peer no estructuradas, los nodos no están organizados de un modo particular, sino que se conectan entre sí de forma aleatoria.

Las redes no estructuradas tienen una serie de ventajas sobre las redes estructuradas. Son más fáciles de construir y mantener y son más robustas ante fallos de nodos. Sin embargo, las redes no estructuradas también tienen algunas desventajas. Pueden ser más lentas que las redes estructuradas, y son más vulnerables a los ataques de denegación de servicio.

La búsqueda de información/datos se realiza de la siguiente manera, cuando un nodo quiere encontrar un recurso en la red, envía una consulta de búsqueda a todos sus vecinos. Estos vecinos reenvían la consulta a su vez a todos sus vecinos, y así sucesivamente. Este proceso continúa hasta que la consulta alcanza todos los nodos de la red.

Este tipo de búsqueda presenta una serie de inconvenientes, genera una gran cantidad de tráfico de señalización en la red, utiliza más CPU y memoria en todos los nodos de la red y no garantiza que las consultas de búsqueda siempre se resuelvan. Además, si un nodo busca datos compartidos en una pequeña cantidad de nodos, es muy poco probable que la búsqueda tenga éxito [95].

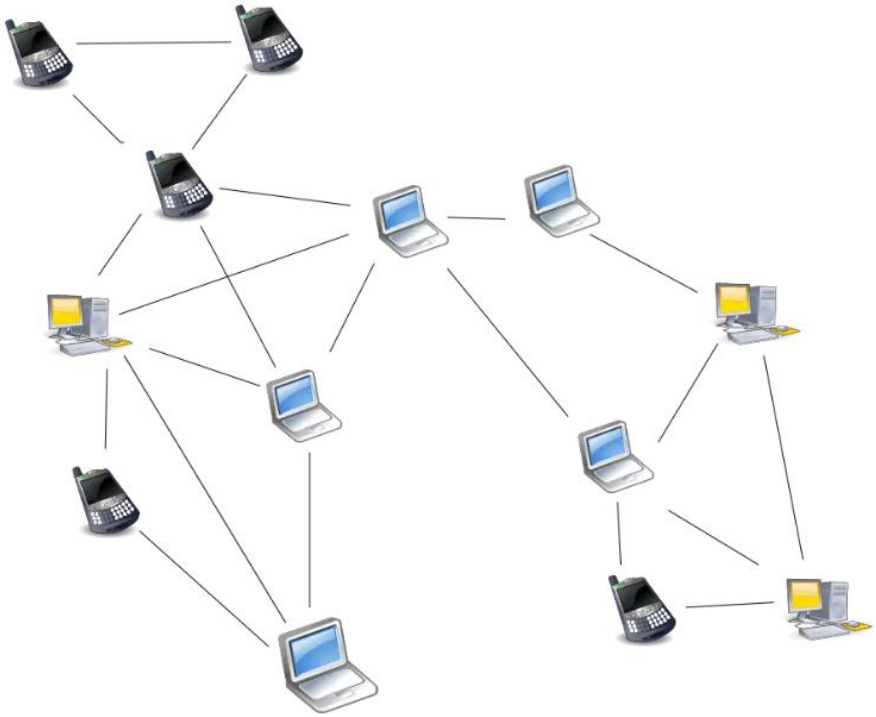


Figura 4-12 Ejemplo Red P2P no estructurada [89]

En las redes peer-to-peer **estructuradas**, los nodos están organizados en una estructura específica. Esta estructura permite a los nodos buscar archivos o recursos de forma rápida y eficiente, incluso si el recurso es raro.

El tipo más común de redes P2P estructuradas implementa una tabla hash distribuida (DHT), en la que se utiliza una variante del hash para asignar la pertenencia de cada archivo a un nodo específico. Esto permite a los nodos buscar recursos en la red utilizando una tabla hash: es decir, los pares (clave, valor) se almacenan en la DHT y cualquier nodo participante puede recuperar de forma eficiente el valor asociado a una clave determinada. Sin embargo, para enrutar el tráfico de forma eficiente a través de la red, los nodos estructurados deben mantener listas de vecinos que satisfagan criterios específicos. Esto los hace menos robustos en redes con una alta tasa de rotación (es decir, con un gran número de nodos que se unen y abandonan la red con frecuencia). Las soluciones basadas en DHT, adolecen de varios problemas, como el alto coste en el descubrimiento de recursos y el balanceo de carga.

Algunas redes que utilizan DHT son Tixati [96], una alternativa a BitTorrent, la red Kad [97], la botnet Storm, YaCy [98]y la red de distribución de contenidos Coral [99]. Proyectos de investigación con DHT son el proyecto Chord, Kademlia, almacenamiento PAST, P-Grid, y CoopNet [100]. Las redes basadas en DHT también se han utilizado para mejorar el descubrimiento de recursos en los sistemas de “Grid computing” o computación en malla[2], ya que ayuda a la gestión de recursos y a la programación de aplicaciones [101].

[2] Grid Computing o computación en malla se refiere a la conexión descentralizada de un grupo de ordenadores con el propósito de crear un superordenador virtual. La esencia de esta solución radica en la optimización del uso de la infraestructura, ya que permite aprovechar eficientemente la capacidad de procesamiento disponible en cada uno de los nodos conectados, permitiendo distribuir la carga de trabajo de manera dinámica según las necesidades en un momento dado.

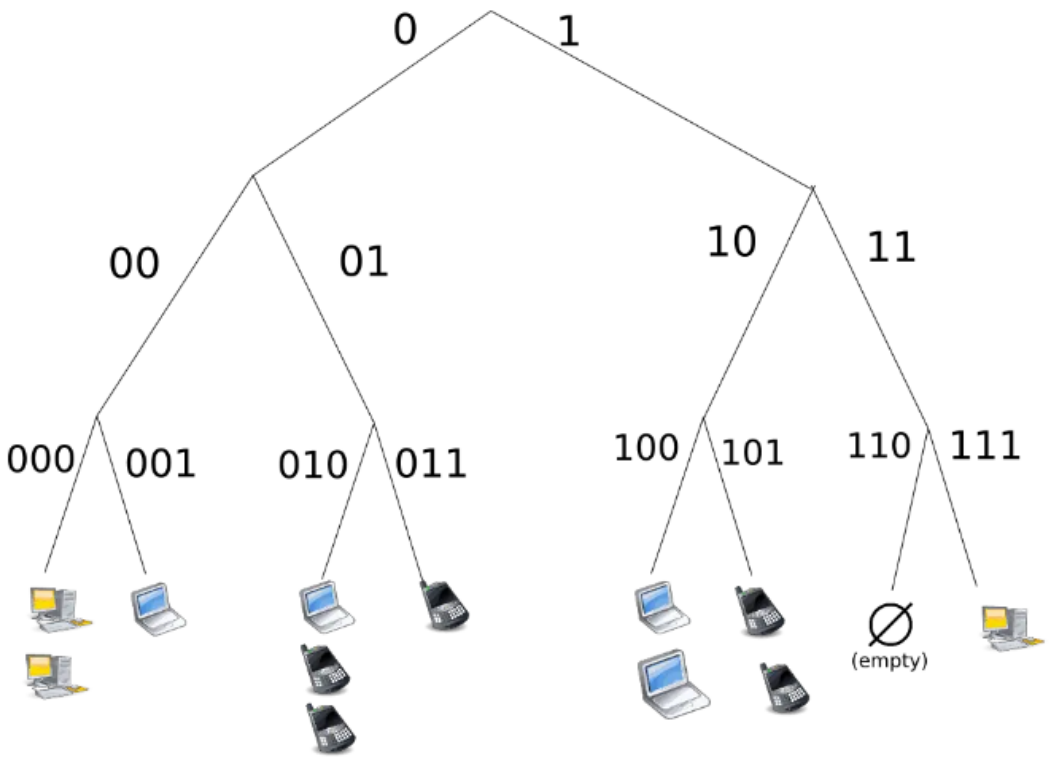


Figura 4-13 Ejemplo Red P2P estructurada [89]

Centralizadas o descentralizadas

Un sistema P2P puede utilizar indexación centralizada, descentralizada, distribuida o híbrida para localizar nodos, recursos compartidos o grupos en una red P2P.

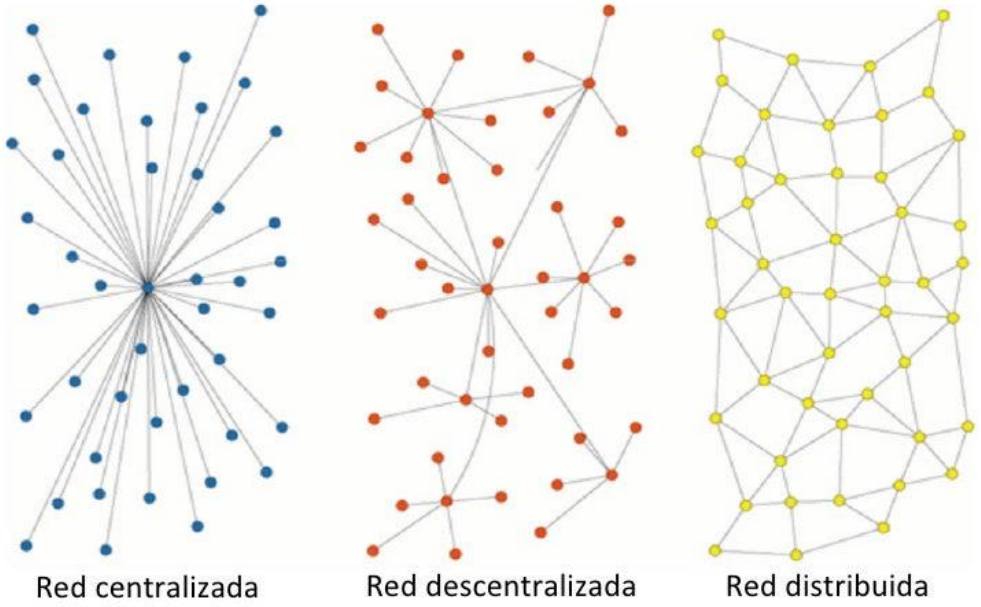


Figura 4-14 Redes P2P por su configuración [102]

En la indexación centralizada, el índice de la red P2P se almacena en uno o varios servidores centralizados que a menudo se denominan rastreadores [103]. Ejemplos comunes de sistemas P2P que utilizan indexación centralizada son Napster [104] y BitTorrent [105].

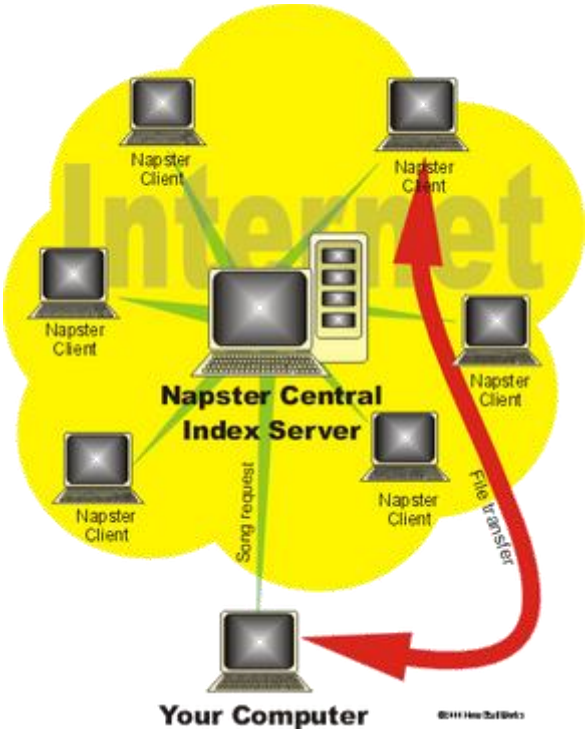


Figura 4-15 Arquitectura Napster [104]

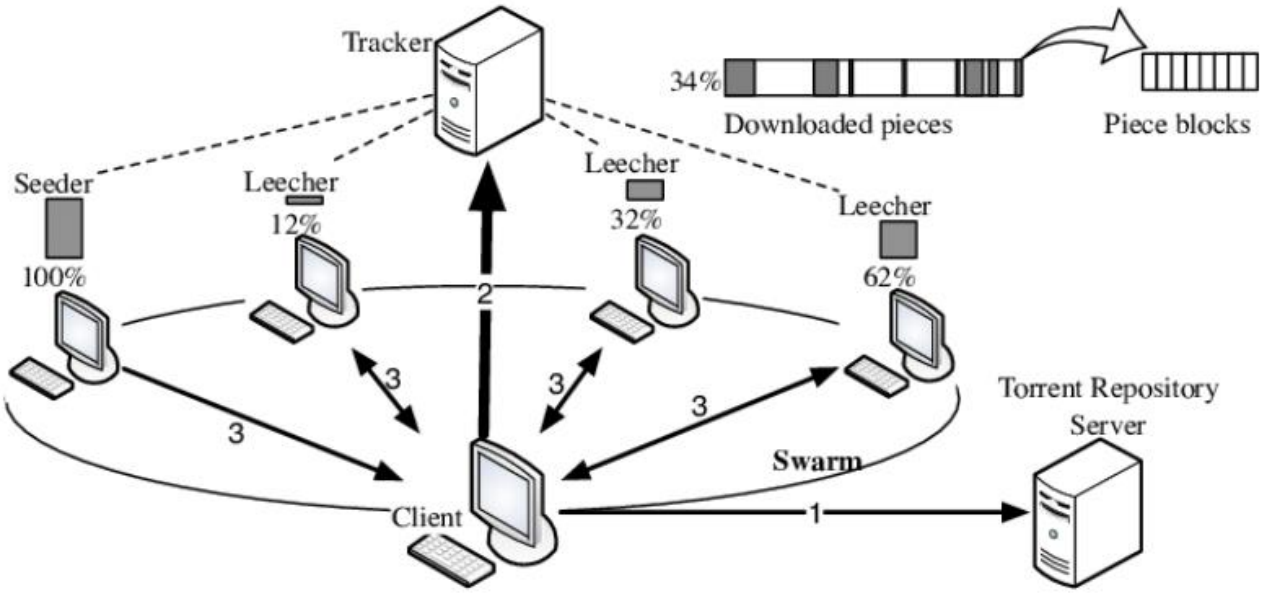


Figura 4-16 Arquitectura BitTorrent [105]

En la indexación distribuida, no hay almacenamiento centralizado para el índice, sino que el índice se distribuye entre los nodos en una red P2P. Por ejemplo, sistemas P2P como Gnutella v0.4 [106] y Freenet [107] dependen de la indexación distribuida.

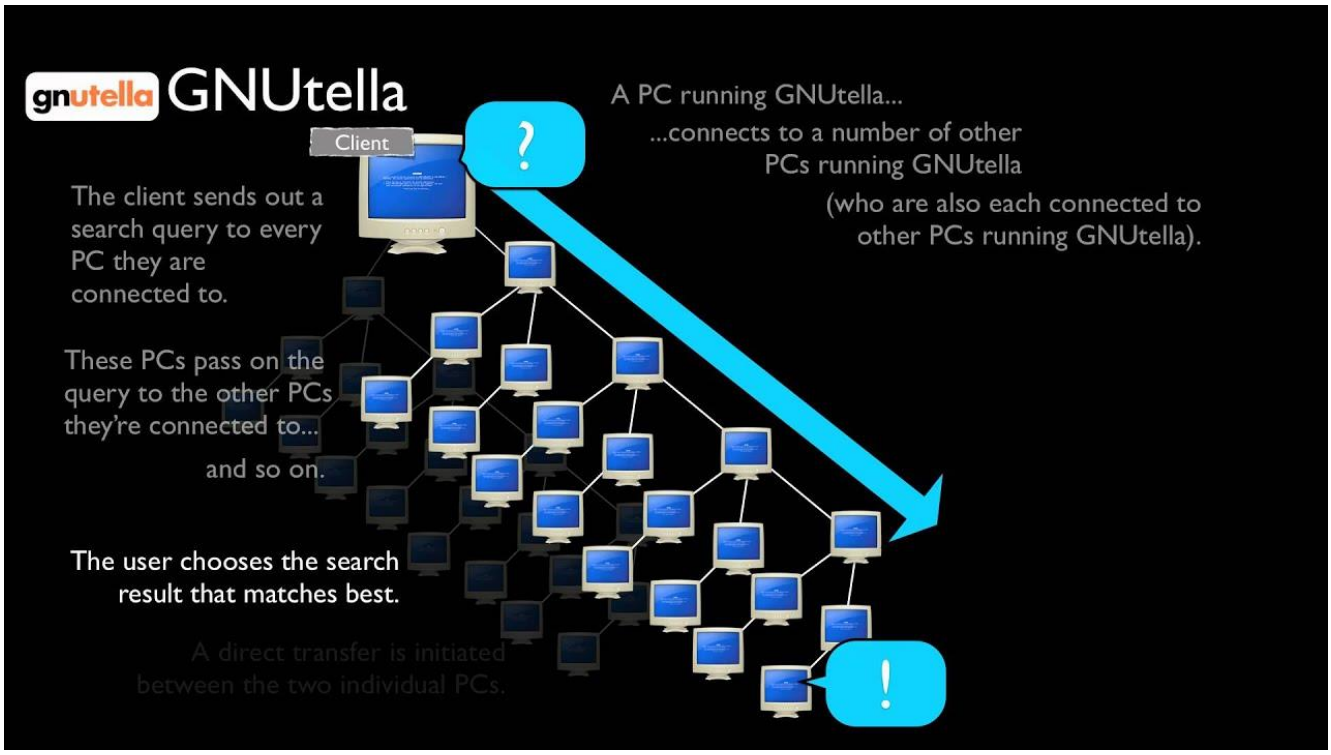


Figura 4-17 Arquitectura GNUtella [108]

En la indexación híbrida, la responsabilidad de mantener el índice de una red P2P se distribuye a un pequeño subconjunto de nodos llamados supernodos. Por lo general, en los sistemas P2P híbridos, los supernodos mantienen los índices de los nodos ordinarios conectados a ellos [109]. En cierto sentido, cada supernodo actúa como un servidor centralizado para un subconjunto de nodos ordinarios.

En la indexación centralizada, los servidores suelen formar parte de la infraestructura del servicio, mientras que en la indexación híbrida, los supernodos suelen ser dispositivos de los usuarios. Ejemplos de sistemas P2P que utilizan indexación híbrida son Skype, Gnutella v0.6 y FastTrack.

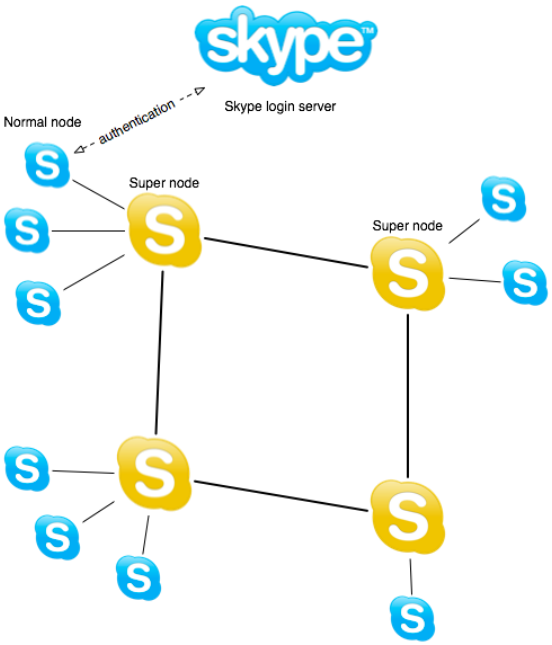


Figura 4-18 Arquitectura Skype [110]

4.3.5 Redes I2P (The Invisible Internet Project)

La red I2P es un tipo de red P2P descentralizada, programada en Java, que proporciona una capa de red de anonimicidad, protegiendo de la censura, vigilancia gubernamental y vigilancia online. Distribuye el tráfico de tal manera que hay una probabilidad infima de que una tercera persona pueda interceptarla [111].

Figura 4-19 Mascota I2P, Itoopie [112]

El Proyecto de Internet Invisible (I2P) comenzó en 2002. La visión del proyecto, según se describe en una entrevista con Lance James, era que la Red I2P "proporcionara pleno anonimato, privacidad y seguridad al más alto nivel posible. Una Internet descentralizada y de igual a igual significa que ya no tendrás que preocuparte por tu ISP controlando tu tráfico. Esto permitirá realizar actividades sin problemas y cambiará la forma en que entendemos la seguridad e incluso Internet, utilizando criptografía de clave pública, esteganografía IP y autenticación de mensajes. La Internet que debería haber sido, será pronto". Desde entonces, I2P ha evolucionado para especificar e implementar un conjunto completo de protocolos de red capaces de proporcionar un alto nivel de privacidad, seguridad y autenticación para una variedad de aplicaciones [113].

La red I2P es una red superpuesta de igual a igual completamente cifrada. Un observador no puede ver el contenido, la fuente o el destino de un mensaje. Nadie puede ver de dónde proviene el tráfico, a dónde se dirige o cuál es su contenido. Además, los transportes de I2P ofrecen resistencia al reconocimiento y bloqueo por parte de censores. Dado que la red depende de pares para enrutar el tráfico, el bloqueo basado en la ubicación es un desafío que aumenta con la red. Cada enrutador en la red participa en hacer que la red sea anónima. Excepto en casos en los que sería inseguro, todos participan en enviar y recibir tráfico de la red.

El software central (Java) incluye un enrutador que introduce y mantiene una conexión con la red. También proporciona aplicaciones y opciones de configuración para personalizar la experiencia y flujo de trabajo.

La red proporciona una capa de aplicación para servicios, aplicaciones y gestión de red. La red también tiene su propio DNS único que permite el autohospedaje y la duplicación de contenido desde Internet (Clearnet). La red I2P funciona de la misma manera que Internet. El software Java incluye un cliente BitTorrent y correo electrónico, así como una plantilla de sitio web estática.

I2P utiliza la criptografía para lograr diversas propiedades para los túneles que construye y las comunicaciones que transporta. Los túneles de I2P utilizan transportes, NTCP2[3] y SSU2[4], para ocultar el tráfico que se transporta sobre ellos. Las conexiones están cifradas de enrutador a enrutador y de cliente a cliente (de extremo a extremo). Se proporciona secreto directo (forward Secrecy) para todas

[3] NTCP2 – Nueva versión de NTCP (TCP basado en NIO (Java NIO (new I/O))
[4] SSU – Secure Semireliable UDP

las conexiones. Debido a que I2P tiene direcciones criptográficas, las direcciones de red de I2P solo pertenecen al usuario que las generó.

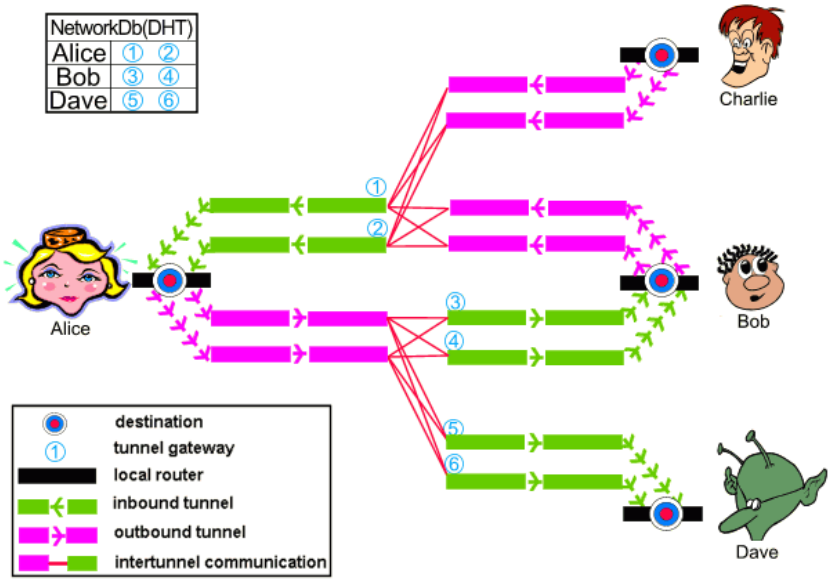


Figura 4-20 Funcionamiento I2P [114]

La red está formada por pares ("enrutadores") y túneles virtuales unidireccionales de entrada o de salida. Los enrutadores se comunican entre sí mediante protocolos construidos sobre mecanismos de transporte existentes (TCP, UDP), transmitiendo mensajes. Las aplicaciones de cliente tienen su propio identificador criptográfico ("Destino") que le permite enviar y recibir mensajes. Estos clientes pueden conectarse a cualquier enrutador y autorizar la asignación temporal ("Lease") de algunos túneles que se utilizarán para enviar y recibir mensajes a través de la red. I2P tiene su propia base de datos de red interna (utilizando una modificación de la DHT de Kademlia) para distribuir información de enrutamiento y contacto de manera segura.

La red I2P funciona prácticamente descentralizada, con excepción de lo que se llaman "Reseed Servers", que almacenan una copia de los datos de la red y garantizar que la red siga siendo accesible. Básicamente, no hay una forma buena y confiable de ejecutar al menos un nodo de inicio permanente que el personal no participante en la red pueda encontrar. Una vez conectado a la red, el enrutador solo descubre miembros de la red construyendo túneles "exploratorios", pero para hacer la conexión inicial, se requiere un "reseed host" para crear conexiones e integrar un nuevo enrutador a la red. Los "reseed servers" pueden observar cuando un nuevo enrutador ha descargado una semilla (seed) de ellos, pero nada más sobre el tráfico en la red I2P.

En la red I2P los proxies a Internet son administrados por voluntarios y son servicios centralizados. Los beneficios de privacidad al participar en la red I2P provienen de permanecer en la red y no acceder a Internet. El navegador Tor o una VPN de confianza son opciones mejores para navegar por Internet de manera privada.

Seguridad en I2P

La red I2P es una red que proporciona un alto grado de seguridad debido a los siguientes factores [111].

- Utiliza cifrado de extremo a extremo y protege los extremos de la ruta por la que viajan los datos convirtiéndolos en identificadores criptográficos, asegurándolos con claves públicas.

- Los túneles unidireccionales de I2P separan el tráfico entrante y saliente, proporcionando mayor privacidad. Gracias a estos enrutadores de proxy de entrada y salida, los remitentes y destinatarios no necesitan revelar sus direcciones IP.
- A diferencia del enrutamiento de cebolla de Tor, I2P utiliza el llamado enrutamiento de ajo (garlic routing). Esto divide tu mensaje en mensajes más pequeños llamados "clavos". Todos estos están cifrados y viajan por separado hasta sus destinos. Como resultado, es casi imposible que un interceptor obtenga todo el mensaje y realice un análisis de tráfico.

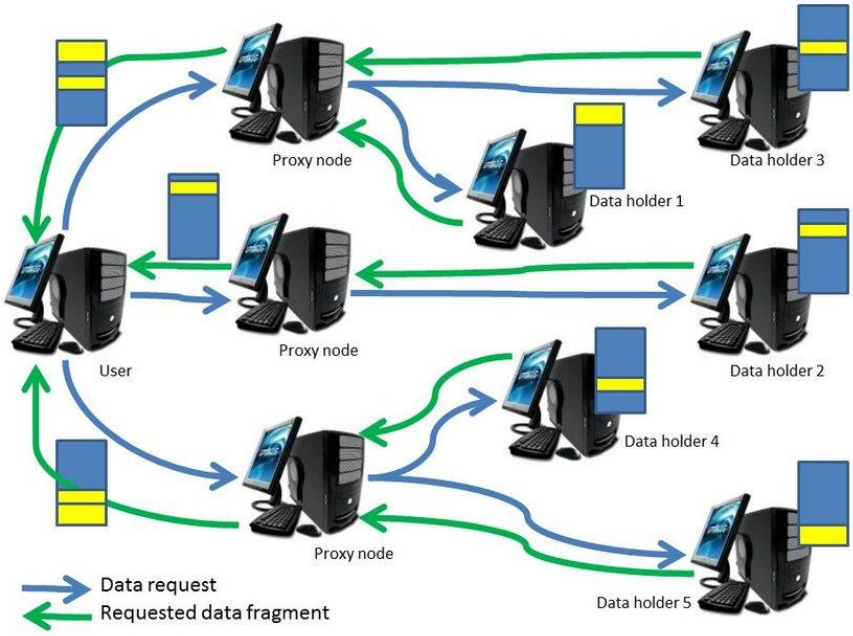


Figura 4-21 Funcionamiento I2P [115]

- El sistema es completamente descentralizado. Como se mencionó anteriormente, I2P se distribuye en miles de dispositivos diferentes y no depende de una entidad única. Si un dispositivo es pirateado, no comprometerá todo el sistema.

4.3.6 Freenet

Freenet es otra plataforma de P2P para la comunicación y publicación resistente a la censura, permite a los usuarios publicar y acceder a información sin temor a la censura o la vigilancia. Freenet utiliza una red descentralizada y anónima para garantizar la privacidad y la seguridad de los usuarios. Es un software de código abierto y está disponible para descargar en Windows, Linux, macOS y Android.

Figura 4-22 Logo de Freenet [116]

Freenet fue creado por Ian Clarke, un científico de la computación que publicó en la red su proyecto de final de carrera en 1999, el ensayo "A Distributed Decentralised Information Storage and Retrieval System". En este ensayo, Clarke describía un algoritmo que, al ser ejecutado por un grupo interconectado de nodos, proporcionaría un sistema robusto de almacenamiento indexado con claves y recuperación sin ningún elemento de control centralizado o administración. Al publicar su ensayo, Clarke invitó a cualquiera que quisiera ayudarle a implementar su diseño, dando como resultado un proyecto de software libre que acabó siendo conocido como Freenet.

Es una red que posibilita el intercambio de archivos de manera anónima. Una vez compartido un archivo o actualizada una página en la Freenet, el documento queda insertado en la red, por lo que una vez finalizada la inserción ya no hace falta que el nodo siga operativo para poder acceder al contenido. Por lo tanto, podría describirse como una especie de caché gigantesco que se distribuye a través de sus nodos. Estos nodos no están jerarquizados y se transmiten mensajes y documentos entre ellos. Los nodos pueden funcionar como nodos finales o intermedios de enrutamiento [117].

Los nodos finales almacenan datos específicos, mientras que los nodos intermedios se encargan de enrutar los datos entre los nodos finales. Cada nodo almacena una tabla de enrutamiento que asocia los nodos con un historial. Esto permite que los datos se encuentren y se transmitan de forma eficiente a través de la red.

La descentralización de Freenet dificulta la censura y el seguimiento de los usuarios. Los datos se almacenan de forma anónima, por lo que no es posible identificar al usuario que los ha almacenado o accedido a ellos.

4.3.7 ZeroNet

ZeroNet es otro proyecto que busca descentralizar la web, pero difiere de Freenet en algunos aspectos. ZeroNet utiliza tecnologías blockchain y redes P2P para crear un Internet más resistente a la censura y la manipulación centralizada.

Figura 4-23 Logo de ZeroNet [118]

ZeroNet posee las siguientes caracteristicas:

- **Descentralización**: Al igual que Freenet, ZeroNet busca descentralizar la web eliminando la necesidad de servidores centralizados. En lugar de depender de servidores web tradicionales, las páginas web de ZeroNet son almacenadas y distribuidas por una red de nodos peer-to-peer.
- **Blockchain:** ZeroNet utiliza la tecnología blockchain para gestionar la estructura de su red y la publicación de contenido. Cada sitio web en ZeroNet tiene su propia dirección Bitcoin[5]

[5] Las direcciones Bitcoin en ZeroNet son direcciones únicas que se utilizan para recibir pagos en Bitcoin en la red ZeroNet. Estas direcciones se generan utilizando el formato estándar de dirección Bitcoin (cadena de 26-35 caracteres alfanuméricos) y se utilizan para facilitar las transacciones en la red.

asociado a un certificado digital, y los usuarios pueden modificar el contenido del sitio al firmar las actualizaciones con la clave privada asociada a esa dirección.

- **Resistencia a la censura:** Debido a su estructura descentralizada y su uso de la tecnología blockchain, ZeroNet es resistente a la censura. No hay un punto central de control que pueda ser atacado o bloqueado para detener la disponibilidad de un sitio web específico.
- **Actualizaciones sin servidor:** La actualización de contenido en ZeroNet se realiza a través de la tecnología blockchain, lo que significa que los sitios web pueden actualizarse sin depender de un servidor central.
- **Seguridad:** El uso de criptografía de Bitcoin por parte de ZeroNet garantiza que los datos y las comunicaciones de los usuarios permanezcan seguros. Esto proporciona protección contra violaciones de datos y acceso no autorizado.
- **Privacidad:** La naturaleza descentralizada de ZeroNet dificulta el seguimiento de la actividad de los usuarios. Esto proporciona un nivel de anonimato que no es posible en las plataformas web tradicionales.
- **Anonimato:** Aunque ZeroNet no está diseñado específicamente para proporcionar anonimato como Freenet, ofrece ciertos niveles de privacidad. La identidad del creador de un sitio web se puede mantener de manera más anónima debido al uso de direcciones Bitcoin.

4.3.8 Otras redes resistentes a la censura

Además de las redes mencionadas anteriormente, existen multitud de redes y tecnologías diseñadas para ser resistentes a la censura y ofrecer mayor privacidad y anonimato. No obstante, ninguna red es completamente invulnerable.

Como ejemplo vamos a nombrar algunas de ellas [119]:

- **GNUnet:** es un marco para la creación de redes peer-to-peer seguras y descentralizadas. Está diseñado para proporcionar privacidad y resistencia a la censura al permitir a los usuarios compartir información de manera anónima.
- **IPFS (InterPlanetary File System):** es un sistema de archivos distribuido diseñado para hacer que la web sea más rápida, segura y abierta. La información se almacena en nodos distribuidos en lugar de servidores centralizados.
- **Matrix:** es una red descentralizada de mensajería que permite la comunicación en tiempo real a través de diferentes servicios y servidores. Puede ser utilizado para crear aplicaciones de mensajería seguras y resistentes a la censura.
- **Mesh Networks:** Las redes en malla son sistemas de comunicación donde cada dispositivo conectado actúa como un nodo y puede enrutar datos para otros dispositivos en la red. Estas redes son resistentes a la censura ya que no dependen de una infraestructura centralizada.
- **OpenBazaar:** es un mercado descentralizado que utiliza la tecnología de la cadena de bloques. Permite a los usuarios comprar y vender bienes y servicios sin la necesidad de intermediarios centralizados.
- **RetroShare:** es una plataforma de comunicación descentralizada que incluye funciones de mensajería, intercambio de archivos y foros. La comunicación entre usuarios se realiza de forma cifrada y se basa en una red de confianza.

- **Diaspora:** es una red social descentralizada que permite a los usuarios tener más control sobre sus datos. Los usuarios pueden unirse a "pods" independientes que actúan como nodos en la red global de Diaspora.
- **WhisperSystems / Signal:** es una aplicación de mensajería que utiliza cifrado de extremo a extremo para proteger la privacidad de las comunicaciones. Es conocida por su enfoque en la seguridad y la privacidad y es resistente a la censura.
- **Tox:** es una plataforma de mensajería instantánea descentralizada que ofrece cifrado de extremo a extremo. Permite a los usuarios comunicarse sin depender de un servidor centralizado.
- **Syndie:** es un sistema para compartir información de manera anónima y resistente a la censura. Utiliza redes de datos distribuidas y cifrado para garantizar la privacidad de los usuarios.
- **Osiris, OneSwarm, Tribler:** son redes anónimas que permiten a los usuarios compartir archivos y comunicarse de forma segura.
- **GlobaLeaks, SecureDrop:** son plataformas de denuncia de filtraciones de información que permiten a los usuarios compartir información de forma segura y anónima.

4.4 Minimizar la huella digital

La huella digital (ya definida en el capítulo introductorio) es el rastro que se deja en Internet cada vez que se utilizan diferentes servicios, aplicaciones o dispositivos. La huella digital puede revelar información personal, preferencias, hábitos, ubicación y otros datos que pueden ser usados por terceros para rastrear, enviar publicidad, violar la privacidad o incluso robarte identidades [120].

Toda huella digital se compone de dos tipos de información: activa y pasiva [121].

- La **parte activa** de la huella digital consiste en la información que se deja deliberadamente en Internet. Por ejemplo, participar en foros, suscribirse a boletines de noticias, etc., todo ello requiere una aportación deliberada del usuario. Esta aportación se integra en una huella digital.
- La **parte pasiva** de una huella digital se crea cuando se recopila información sobre un usuario sin su conocimiento, lo cual ocurre todo el tiempo. En cuanto se visita un sitio web por primera vez, ese sitio web reconoce al individuo como usuario (puede identificarlo en función de su huella digital). Cada vez que se visita ese sitio web a partir de entonces, se recopila más información sobre la persona, como las páginas que se visitan, los enlaces en los que se hace clic, el tiempo se pasa en un sitio, etc. Gran parte de esta información se recopila mediante el uso de cookies, que son fragmentos de código que se almacenan en el dispositivo con el propósito de recopilar datos sobre un individuo. Irónicamente, la mayoría de los sitios web solicitan que se acepten las cookies, lo que aumenta la huella digital activa (aunque la recopilación de datos realizada por las cookies siga siendo pasiva).

4.4.1 Catálogo de buenas prácticas

Con el conocimiento de que la seguridad al 100% no existe y el uso de Internet crea indefectiblemente una huella digital, vamos a describir una serie de acciones o buenas prácticas para disminuir el rastro que se genera en Internet.

CONFIGURACIÓN

- **Uso de sistemas operativos "live-USB"** o en vivo son aquellos que no tienen que estar instalados en el ordenador y se pueden arrancar y ejecutar desde un USB, DVD o CD.

Durante el arranque el SO se carga en la RAM trabajando desde ahí [122] (por ejemplo TAILS). Debido a esta forma de funcionamiento, estos SOs no dejan rastro en los ordenadores.

- **Uso de un sistema operativo seguro:** ciertos sistemas operativos están especialmente diseñados para proporcionar privacidad y seguridad (HarmonyOS, KasperskyOS).
- **Actualizaciones automáticas:** es recomendable tener activada la función de actualizaciones automáticas de nuestros sistemas, especialmente los personales, ya nos proporcionan [123]:
    - Corrección de vulnerabilidad publicadas: mediante la instalación de los parches de seguridad.
    - Solución de errores: Todo desarrollo SW tiene errores (bugs) que se van descubriendo a medida que se utiliza. Dichos bugs pueden producir comportamientos inesperados de las herramientas.
    - Nuevas funcionalidades
- **Actualizar el software regularmente:** contar con la versión más reciente del software en los ordenadores y dispositivos móviles es crucial para salvaguardar la información. Esto implica no solo actualizar el antivirus, sino también todos los programas instalados. Al mantener todo al día, se dificulta el acceso de los hackers a los sistemas.
- **Configuración de privacidad y seguridad**: es conveniente revisar de las cuentas en línea, como redes sociales, correo electrónico, servicios de nube, etc. Ajustar los permisos que se otorgan a cada sitio o aplicación, y limitar lo que pueden ver o acceder otros usuarios. Por ejemplo, se puede desactivar la geolocalización, el historial de búsqueda, las notificaciones, el acceso a la cámara o micrófono, etc, limitando la información personal que se comparte públicamente.
    - **Desactivar JavaScript:** aunque algunos sitios web no funcionan bien sin JavaScript, este lenguaje de programación sirve a los anunciantes para rastrear.
    - **Cifrado de Archivos:** utilizar herramientas de cifrado para proteger los archivos y documentos sensibles antes de cargarlos a la nube o enviarlos por correo electrónico.
- **Bloqueo de Cookies y Rastreadores:** configurar el navegador para bloquear cookies y rastreadores de terceros y eliminar cookies regularmente para evitar el seguimiento en línea.
- **Software de Bloqueo de Anuncios y Rastreadores:** utilizar software de bloqueo de anuncios y rastreadores para evitar que los sitios web recopilen datos sobre el comportamiento en línea.
- **Control y Bloqueo de Ventanas Emergentes y Plugins:** las ventanas emergentes con publicidad que aparecen mientras se navega en Internet puede conducir a situaciones de riesgo. Algunas webs intentan instalar extensiones en los ordenadores de forma subrepticia. Para ello, utilizan técnicas de ingeniería social para engañar y y proceder a la instalación. Estas extensiones pueden recopilar datos de navegación, mostrar publicidad, minar criptomonedas o difundir malware. Es importante bloquear este tipo de complementos para proteger los dispositivos electrónicos [123].
- **Uso de Redes Privadas Virtuales (VPNs):** el uso de estas redes sirve para ocultar la dirección IP y la ubicación.

- Cambiar la dirección MAC de los dispositivos[6]
- **Antivirus**: utilizar un antivirus que emplee Inteligencia Artificial para detectar malware y ataques de día cero y que verifique las actualizaciones del software [124].

NAVEGACIÓN

- **Navegación segura mediate HTTPS:** HTTPS (Protocolo de Transferencia de Hipertexto Seguro) es una versión segura de HTTP (Protocolo de Transferencia de Hipertexto). Utilizar un protocolo criptográfico llamado Transport Layer Security (TLS) para cifrar la comunicación entre el cliente y el servidor, garantiza que los datos se transmitan de forma segura y no puedan ser interceptados por terceros.

  HTTP es un protocolo sin cifrar, lo que significa que los datos transmitidos entre el cliente y el servidor están en texto plano. Esto hace vulnerable al protocolo a la escucha y a los ataques activos o de tampering.

  En una sesión de navegación segura, en la barra de navegación, aparecen las siglas https.

Figura 4-24 – Identificación de HTTP y HTTPS [125]

- **Usar la barra de navegación para escribir las direcciones:** por lo general, se buscan las páginas web de uso diario en los motores de búsqueda. Sin embargo, es importante evitar hacerlo, ya que los piratas informáticos pueden crear clones de estas páginas y colocarlos entre los primeros resultados [126].
- **No utilizar equipos compartidos:** los ordenadores públicos son vulnerables a ataques informáticos. Por eso, es mejor no iniciar sesión en servicios privados en ellos. Si hay que hacerlo, asegurarse de cerrar sesión inmediatamente y de cambiar las contraseñas después.
- **Motores de Búsqueda Privados:** utilizar motores de búsqueda que no rastrean las consultas de búsqueda ni almacenen información personal, como DuckDuckGo.
- **Redes WiFi Seguras:** evitar usar redes WiFi públicas y abiertas. Para usar una red WiFi pública, considerar el uso de una VPN para cifrar la conexión. No hay forma de saber si una conexión WiFi es segura y cualquiera podría estar espiando. Se pueden visitar sitios genéricos, pero hay que evitar enviar información privada [127].
- **Limitar Compartir en Dispositivos:** limitar la información personal compartida a través de dispositivos como teléfonos inteligentes y dispositivos inteligentes para el hogar.
- **Activar el modo oculto /navegación privada**: el modo oculto o incógnito de los navegadores, posibilita el uso de Internet sin guardar información en tu computadora [128].

[6] Dirección MAC (Media Access Control) es un identificador único de 48 bits que se asigna a cada tarjeta de red de un dispositivo. Se utiliza para identificar de forma única a un dispositivo en una red.

- **Eliminar sus datos de forma regular:** esto incluye el historial de navegación, cookies y archivos temporales.

CONTRASEÑAS

- **Contraseñas complejas:** para proteger la seguridad, es conveniente el uso de contraseñas complejas y que se cambien con frecuencia. Una contraseña segura es larga (tiene al menos 12 caracteres e idealmente más) y contiene una combinación de letras mayúsculas y minúsculas, además de símbolos y números. Mientras más compleja y dedicada sea una contraseña, más difícil será descifrarla. Si se necesita apuntarlas, guardarlas en un sitio que solo el individuo conozca.
- **Administrador de contraseñas:** usar un administrador de contraseñas puede ayudar a generar, almacenar y administrar todas las contraseñas en una única cuenta segura en línea.
- **Mantener la privacidad de las contraseñas:** evitar compartir las contraseñas con otros o escribirlas en papel.
- **Autenticación de Dos Factores (2FA):** habilitar la autenticación de dos factores siempre que sea posible para agregar una capa adicional de seguridad a las cuentas en línea.
- Evitar usar la misma contraseña para todas las cuentas y cambiarlas con frecuencia.
- Evitar almacenar contraseñas en los gestores de contraseñas de los navegadores:
    - En el caso de acceso ilícito a un dispositivo, se podrá acceder a los servicios que tengan las credenciales almacenadas. El atacante solo tendrá que abrir el servicio en cuestión en el navegador y la función de autocompletado hará el resto.

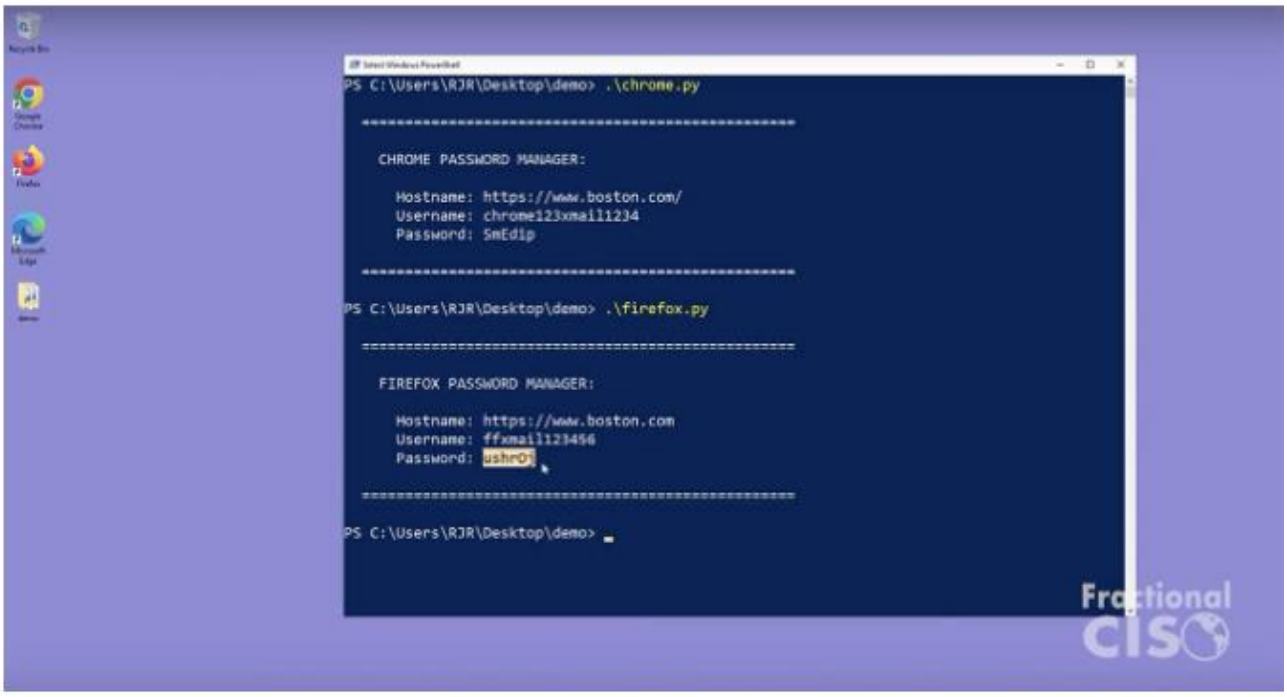


Figura 4-25 Demostración de cómo extraer contraseñas almacenadas en Navegadores Google Chrome, Mozilla Firefox y Microsoft Edge [129].

    - La forma en que se gestionan las contraseñas depende del navegador y del sistema operativo que se use. En algunos casos, es necesario conocer la contraseña de usuario del sistema o una contraseña maestra para poder acceder a las contraseñas almacenadas. En otros casos, no es necesario nada, por lo que un tercero podría ver

todas las contraseñas almacenadas. Si las contraseñas están protegidas con una clave, el atacante podría evadir esta medida de seguridad.

- En Internet hay diferentes páginas web que describen el proceso a realizar para obtener las credenciales de un navegador [130] [131] [132] [133].

- **No registrase con cuentas de Google, Facebook o Twitter:** ciertas páginas dan la opción de registrase o crear cuentas a usando redes sociales. Esto se realiza a través del protocolo OAuth (estándar abierto de autenticación), ya que las grandes empresas comparten alguna información de sus cuentas con aplicaciones o sitios web de terceros [134].

Figura 4-26 Protocolo OAuth [134]

Esta práctica conlleva que si la cuenta con la que nos hemos registrado (Facebook, Google, etc) es comprometida, todas las cuentas de los sitios en los que hemos usado dicha cuenta, pueden ser también comprometidos.

La página web en la que un individuo se registra tiene acceso a esa información, como correo electrónico, perfil público, teléfono, información de contactos, o incluso poder de publicación en ese perfil.

- **No iniciar sesión con Facebook:** iniciar sesión en sitios web y aplicaciones usando Facebook es una forma rápida y sencilla de hacerlo. Sin embargo, es importante tener en cuenta que, al hacerlo, se da permiso a esa empresa para que acceda a los datos de usuario de Facebook. Esta práctica podría poner en riesgo la información personal.

Figura 4-27 Extracto política de privacidad de META [135].

CORREO ELECTRÓNICO Y SUSCRIPCIONES

- **Correo Electrónico Seguro:** usar servicios de correo electrónico cifrados que ofrecen cifrado de extremo a extremo para proteger los correos electrónicos de miradas indiscretas.

- **Eliminar cuentas de correo electrónico antiguas:** las cuentas de correo electrónico inactivas son un objetivo atractivo para los hackers. Si acceden a una, pueden ver los contactos, hacerse pasar por uno, buscar información personal y probar la misma contraseña en otras cuentas. Si no se revisa el correo electrónico con regularidad, es posible que uno desconozca si ha habido una intrusión hasta que sea demasiado tarde.
- **Eliminar Cuentas No Utilizadas:** cerrar o eliminar cuentas en línea que no son utilizadas con la intención de reducir la cantidad de información disponible en la web.
- **Crear una cuenta de correo electrónico spam:** crear una cuenta específica para beneficiarse de promociones, como descuentos en tiendas, y limitar cualquier información de identificación que se almacene en esa cuenta de correo electrónico. De esa manera, si se filtra y publica una base de datos de un comercio y dicha cuenta se ve comprometida, se puede simplemente borrarla para intentar minimizar la pérdida de información digital. (Ventaja: esta práctica ayuda a reducir drásticamente la cantidad de spam en el correo electrónico).
- **Usar cuentas de correo electrónico o registros temporales:** cuando se quiera acceder a un sitio web que no se conozca o que se desconfíe. Así se evita dar información real y recibir spam o correos maliciosos. Hay servicios que ofrecen crear direcciones de correo electrónico que duran unos minutos o unas horas, como GuerrillaMail, 10 Minute Mail o Temp Mail.
- **Usar correo electrónico y comunicación anónimos**: hay servicios que permiten enviar correos electrónicos y mensajes sin revelar la identidad (ej. ProtonMail, 5ymail, AnonEmail, Tutanota), o bien mediante la creación de cuentas de correo permanentes con datos falsos [136].
- **Usar cuentas de correo "Alias":** para acceder a servicios que exijan una dirección de correo electrónico que exista [137].

  Las cuentas de correo alias son direcciones de correo electrónico que se reenvían a una dirección de correo electrónico principal. Esto significa que cualquier correo electrónico enviado a un alias se entregará a la dirección principal.

  Las cuentas de correo alias protegen la privacidad y seguridad. Por ejemplo, se puede usar un alias para registrarse en un servicio que no se conoce o en el que no se confía. Si el servicio se ve comprometido, los atacantes solo tendrán acceso al alias, no a la dirección de correo electrónico principal del usuario.
- **Cancelar las suscripciones a listas de correo:**
  - Si se ha proporcionado el correo electrónico para obtener un cupón o descuento, es posible que se reciban correos electrónicos no deseados de esas compañías. Se puede anular la suscripción a estas listas de correo para mantener la cuenta de correo más limpia y evitar que terceros recopilen información.
  - Hacer clic en los enlaces de cancelación de suscripción cuando estén incluidos en correos electrónicos legítimos de proveedores legítimos, pero no si el correo electrónico parece ser de un vendedor de spam o un estafador [138].

<u>REDES SOCIALES</u>

- **Limitar las cuentas en las redes sociales**: para muchas personas, desactivar las cuentas de redes sociales es difícil o innecesario.

En lugar de borrarlas, se puede minimizar la exposición y los datos que se comparten con los demás:

- Revisar las configuraciones de privacidad con regularidad.
- Limitar las publicaciones a "solo amigos".
- Desactivar la recopilación de datos de ubicación.
- Eliminar los perfiles de los resultados en buscadores públicos.
- Desactivar las cuentas activas no utilizadas.

- **Evitar compartir información personal:** como fecha de nacimiento, lugar de nacimiento, ocupación y estado civil.
- **No Revelar Información Personal:** hay que ser cauteloso al compartir información personal en línea, incluso en redes sociales. La información que se comparte en línea podría ser utilizada por técnicas de rastreo.
- **Limitar la cantidad de información que comparte sobre la vida personal**: esto incluye fotos, videos y publicaciones.
- **Ser cuidadoso con los "me gusta" y comentarios:** esto puede ser utilizado por los anunciantes para crear un perfil de los intereses de una persona.
- **Desactivar las notificaciones:** esto puede ayudar a reducir la cantidad de datos que se recopilan sobre una persona.
- **No rellenar encuestas:** evitar cuestionarios clickbait en sitios web al azar. Es simplemente otra forma que las compañías utilizan para recopilar información digital.
- **Mensajería Instantánea Segura:** utilizar aplicaciones de mensajería que ofrecen cifrado de extremo a extremo, como Signal o WhatsApp (en modo de conversación secreta).

CONCIENCIACIÓN

- **Educación Digital:** aprender sobre las prácticas seguras en línea y enseñar a familiares y amigos cómo proteger la privacidad.
- **Phishing:** aprender a reconocer correos electrónicos de phishing y sitios web falsos para evitar compartir información personal con estafadores en línea.
- **Controlar la descarga e instalación de programas y aplicaciones:** los programas y aplicaciones que se descargan e instalan pueden recopilar datos sobre las actividades que se realizan en línea.
- **Actualizaciones de Privacidad:** mantenerse informado sobre las políticas de privacidad y los cambios en las configuraciones de privacidad de las plataformas en línea que se utilizan.
- **Actuar rápido después de una filtración:** ante la sospecha de un compromiso de datos en una filtración, actuar de inmediato comunicándolo al banco o proveedor de tarjetas de crédito, informando de la filtración y cambiando las contraseñas que pudieron haber quedado expuestas. Si la contraseña se ha usado para otras cuentas, actualizarlas en todas ellas [139].
- **Abuso de plataformas de confianza:** los hackers utilizan plataformas y herramientas de confianza para ocultar sus actividades. Esto les permite mezclarse con el tráfico normal, lo que dificulta que los analistas humanos y las máquinas los detecten.

TELÉFONO MÓVIL

- **Controlar las aplicaciones de tu teléfono**. listar las aplicaciones del teléfono y revisar los términos y condiciones cuando se obtengan nuevas aplicaciones.

  Numerosas aplicaciones indican el tipo de información que recopilan y su uso.

- **Restringir los permisos de las aplicaciones**: las aplicaciones pueden extraer datos personales del correo electrónico, ubicación, cámara, compartir información, etc.

  Entender cómo se comparte la información con otros sitios y compañías puede ayudar a controlar la huella digital.

- **Eliminar las aplicaciones que no se utilicen:** reduciendo la probabilidad de recopilar información sobre el comportamiento digital.

- **Revisar el uso de dispositivos móviles:** establecer un código de acceso para el dispositivo móvil, de forma que otras personas no puedan acceder en caso de pérdida.

4.5 Hackers – El arte de la ocultación.

Ser indetectable, intocable e ingobernable es una parte clave del mito y el misticismo de los 'hackers'. Sin duda, la capacidad de ocultarse siempre ha sido una parte central del modus operandi de los hackers tanto en el mundo físico como en el digital [140].

Los piratas informáticos (hackers) utilizan una variedad de técnicas para ocultar sus actividades y evadir la detección, lo que dificulta que los atrapen, ya sea la policía o los profesionales de la seguridad.

Los hackers encuentran continuamente nuevas formas más eficientes de infiltrarse en los sistemas, ya sea comprando un exploit confeccionado en la web oscura, innovando nuevas fallas de seguridad o utilizando modelos de lenguaje de inteligencia artificial, como ChatGPT, en ataques de phishing. Sin embargo, también vemos a los hackers utilizar los mismos métodos una y otra vez para irrumpir en sistemas que carecen de seguridad básica. Por lo tanto, si bien los ciberdelincuentes utilizan técnicas de piratería sofisticadas, primero elegirán el camino de menor resistencia.

En otras palabras, los hackers siempre están buscando nuevas formas de atacar los sistemas, pero también son conscientes de que muchos sistemas no tienen implementadas medidas de seguridad básicas. Por lo tanto, los piratas informáticos a menudo intentarán atacar primero estos sistemas más vulnerables, antes de pasar a métodos más sofisticados.

Para ello, los hackers se sirven de diferentes técnicas:

- Cifrado: para ocultar información maliciosa en archivos aparentemente benignos.
- Esteganografía: ocultar información dentro de archivos de imagen u otros archivos.
- Ofuscación: es el proceso de hacer que el código fuente o el código máquina sea más difícil de entender para los humanos o las máquinas. La ofuscación se utiliza a menudo para proteger el código fuente de la ingeniería inversa[7].
- VPNs: Las redes privadas virtuales (VPN) cifran el tráfico de un usuario y lo dirigen a través de un servidor remoto. Los hackers pueden utilizar las VPN para ocultar su dirección IP y su ubicación, así como para acceder a contenido bloqueado.

[7] Es el proceso de descomponer el código fuente para entender cómo funciona.

- Proxies: son servidores que actúan como intermediarios entre el ordenador de un usuario y Internet. Los piratas informáticos pueden utilizar proxies para ocultar su dirección IP y su ubicación, lo que dificulta su rastreo.
- Suplantación de direcciones MAC.
- Uso de máquinas virtuales y Sistemas Operativos "live", de tal manera que cada vez que inician una nueva sesión es como un SO recién instalado.
- Evitar motores de búsqueda que rastrean.
- Uso de botnets[8].

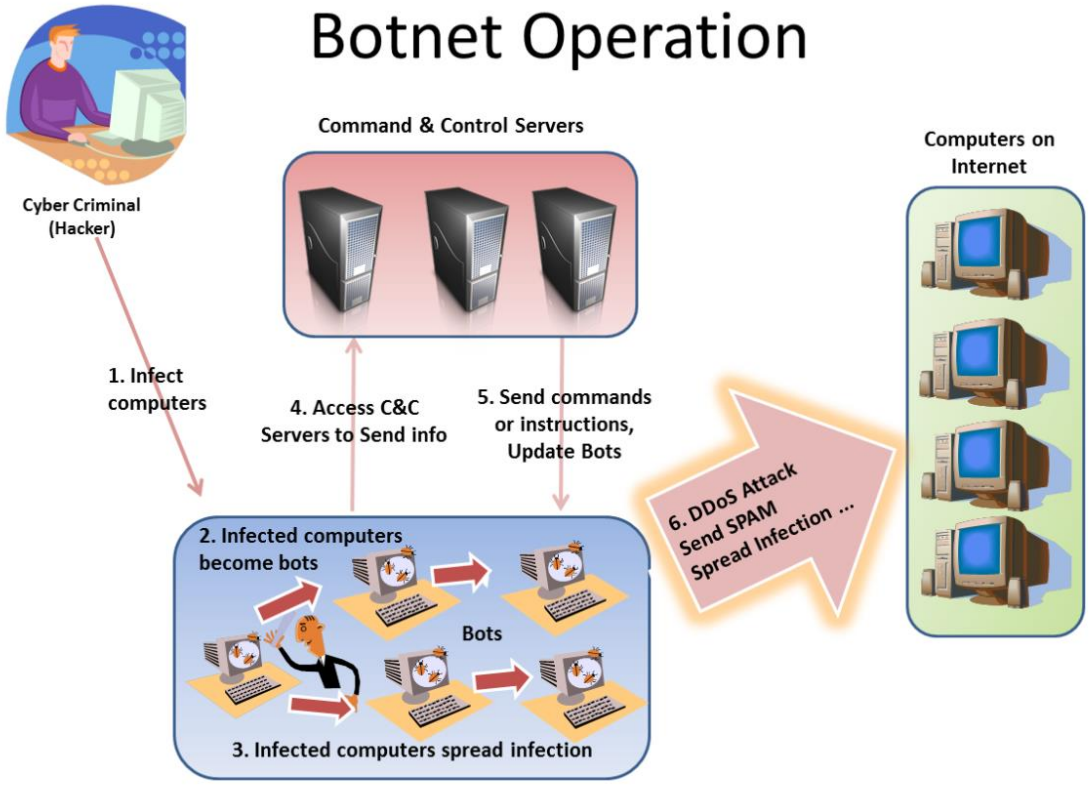


Figura 4-28 Esquema de operación de un Botnet [141].

- Uso de redes anónimas (ACN): Las redes anónimas enrutan el tráfico de un usuario a través de una serie de nodos, lo que dificulta el seguimiento de su dirección IP y ubicación. Los hackers pueden utilizar redes anónimas para ocultar su identidad y ubicación cuando realizan actividades ilegales.

[8] Botnet: son un conjunto de máquinas infectadas que trabajan en coordinación. Un botnet consta [198] de i) varios bots, ii) un servidor de comando y control (C&C) y iii) un botmaster.

- Bots (también llamados zombis [199]): son ordenadores infectados por malware que pueden ser controlados remotamente por un atacante.
- Servidor C&C: El servidor C&C (Command and Control) es el ordenador central que controla el botnet. El servidor C&C envía instrucciones a los bots y recibe información de ellos.
- Botmaster: El botmaster es el atacante que controla el botnet. El botmaster utiliza el servidor C&C para enviar instrucciones a los bots y recibir información de ellos.

- Ingeniería social: es la práctica de manipular a las personas para que revelen información confidencial o realicen acciones que comprometen la seguridad. Los hackers pueden utilizar técnicas de ingeniería social para engañar a las personas para que den sus contraseñas, hagan clic en enlaces maliciosos o abran archivos adjuntos infectados.
- Imitar patrones de vida dentro de la red y se comportarse usuarios legítimos.

Una vez que los hackers han obtenido acceso a un sistema, pueden permanecer en la red durante meses o incluso años utilizando un enfoque sigiloso para evitar la detección del software de escaneo y monitoreo. Por ejemplo, los piratas informáticos analizarán e imitarán el comportamiento de los usuarios autorizados, como solo explorar la red durante el horario laboral normal. Los piratas informáticos también intentarán mezclar su actividad con conexiones de red y protocolos comunes utilizando puertos del sistema de nombres de dominio para enrutar la actividad fraudulenta, disfrazada de consultas aparentemente inofensivas entre redes públicas y privadas. En el caso del fraude por correo electrónico empresarial (BEC), este acceso continuo a un sistema puede ser muy útil para exfiltrar datos [140].

En numerosas ocasiones los actores maliciosos que se esconden dentro de una organización ya son usuarios autorizados dentro de su perímetro[9], lo que les facilita actuar de forma desleal con información y credenciales sensibles. Empleados descontentos, o que han sido sobornados o chantajeados por hackers de fuera de la empresa, podrían verse tentados a filtrar información sensible por beneficio personal o financiero. Esta es, posiblemente, una de las formas más insidiosas en que los hackers se ocultan.

- Uso de complementos de bloqueo de huellas digitales del navegador: Los hackers también pueden utilizar complementos que bloquean las huellas digitales del navegador para evitar ser rastreados.
- Uso de lenguajes de programación poco comunes para codificar malware: Algunos hackers codifican malware en lenguajes de programación poco comunes para evitar la detección.
- Uso de malware y exploits sofisticados.

Los hackers están constantemente desarrollando nuevas formas de ocultar sus actividades y evadir la detección.

4.6 Sistemas Operativos orientados a la seguridad informática

Los sistemas operativos orientados a la seguridad informática se diseñan con un enfoque centrado en proteger la integridad, confidencialidad y disponibilidad de la información. Estos sistemas operativos implementan características y medidas de seguridad avanzadas para proporcionar un alto nivel de seguridad a los datos y sistemas informáticos mitigando riesgos y previniendo ataques cibernéticos.

Actualmente, los sistemas operativos que más utilizamos en la actualidad no son aquellos que se instalan en los ordenadores personales, sino que además tenemos todos los que se instalan en diferentes dispositivos como móviles o tabletas.

[9] Los conocidos como **insiders**, los cuales son personas que tienen acceso legítimo a información privilegiada sobre una empresa u organización. Los insiders pueden ser maliciosos, si actúan con intención de causar daño o obtener un beneficio ilícito, o negligentes, si actúan por error o desconocimiento [200].

Figura 4-29 Sistemas operativos más utilizados en 2023 [142].

En cualquier caso y dentro de los sistemas operativos de los ordenadores personales, podemos mencionar, que aunque no es el más extendendido, el sistema operativo Linux destaca como el sistema operativo más seguro achacables a diversas razones [143]:

- **Requiere conocimientos informáticos**: Su utilización efectiva suele demandar conocimientos informáticos, lo que contribuye a un uso más seguro al permitir que los usuarios tomen decisiones informadas sobre la configuración y el mantenimiento del sistema.
- **Código abierto**: A pesar de ser de código abierto, lo que podría parecer contradictorio en términos de seguridad, Linux se beneficia de esta característica. La comunidad global tiene acceso al código fuente, lo que permite una revisión constante en busca de posibles vulnerabilidades. La sorprendente colaboración de voluntarios de todo el mundo garantiza que cualquier fallo detectado sea abordado de manera rápida y eficiente.
- **Diversidad de distribuciones:** Linux se distingue por la diversidad de distribuciones disponibles, y lo interesante es que no todas representan versiones más recientes. Coexisten numerosas versiones de Linux, cada una con su propio enfoque y características, todas igualmente actualizadas y avanzadas. Esta variedad proporciona opciones adaptadas a diversas necesidades y preferencias de los usuarios.

Como hemos comentado, las distribuciones Linux, priman en seguridad. De esta manera y a modo de ejemplo de sistemas operativos con un enfoque fuerte en seguridad podemos mencionar los siguientes:

- **Qubes OS:** Qubes OS se basa en la idea de la virtualización para crear "dominios" aislados que ejecutan diferentes actividades. Cada dominio tiene un propósito específico, y esto ayuda a prevenir la propagación de amenazas.

- **Tails (The Amnesic Incognito Live System): Tails** es un sistema operativo basado en Linux diseñado para preservar la privacidad y anonimato. Se ejecuta desde un USB o DVD y dirige todo el tráfico a través de la red Tor para ocultar la ubicación del usuario.
- **HardenedBSD:** HardenedBSD es una bifurcación de FreeBSD que se centra en mejorar la seguridad. Incluye características como ASLR (Address Space Layout Randomization), W^X (Write XOR Execute), y otras técnicas para hacer más difícil la explotación de vulnerabilidades.
- **SELinux (Security-Enhanced Linux):** SELinux no es un sistema operativo en sí, sino una extensión del kernel Linux. Proporciona un marco de seguridad avanzado que implementa controles de acceso obligatorios para reforzar la seguridad del sistema.
- **OpenBSD:** OpenBSD es un sistema operativo de código abierto que se enfoca en la seguridad y la calidad del código. Los desarrolladores de OpenBSD revisan y auditan el código constantemente para identificar y corregir posibles vulnerabilidades.
- **Whonix:** Whonix es un sistema operativo diseñado para ser ejecutado en una máquina virtual. Su objetivo principal es proporcionar un entorno seguro y anónimo para la navegación en Internet a través de la red Tor.
- **TOMOYO Linux:** TOMOYO Linux es un módulo del kernel que implementa un control de acceso obligatorio para mejorar la seguridad del sistema Linux. Permite a los usuarios definir políticas detalladas sobre qué acciones pueden realizar los programas.

4.7 Navegadores web – Internet Browsers

Un navegador web es una aplicación de software que permite a los usuarios acceder e interactuar con sitios web. Los navegadores web se utilizan para recuperar y mostrar páginas web, que están formateadas utilizando HTML, CSS y JavaScript entre otros.

Los navegadores web desempeñan un papel importante en el ecosistema de Internet, y están en constante evolución para admitir nuevas funciones y tecnologías. Por ejemplo, los navegadores web modernos admiten conexiones HTTPS seguras, que cifran el tráfico entre el navegador y el servidor web. Esto ayuda a proteger a los usuarios de la escucha y otros ataques.

Los navegadores web también admiten una variedad de otras funciones, como:

- Extensiones: Las extensiones son pequeños programas que se pueden agregar a los navegadores web para mejorar su funcionalidad. Por ejemplo, hay extensiones que pueden bloquear anuncios, agregar nuevas funciones a sitios web de redes sociales y más.
- Marcadores: Los marcadores permiten a los usuarios guardar sus sitios web favoritos para acceder a ellos más tarde.
- Historial: La función de historial permite a los usuarios ver los sitios web que han visitado recientemente.
- Cookies: Las cookies son pequeños archivos de texto que los sitios web almacenan en el ordenador del usuario. Las cookies se pueden utilizar para rastrear la actividad de navegación del usuario, recordar la información de inicio de sesión y más.

Se tiene a disposición de los usuarios una amplia cantidad de navegadores. Cada navegador se ha diseñado con unas características específicas centrándose en la privacidad, motivos organizativos, seguridad, etc [144]. Dependiendo de las necesidades y preferencias de los usuarios, unas funciones tendrán mayor prioridad que otras, por eso la elección de un navegador es una decisión personal.

Figura 4-30 Muestra de la diversidad de navegadores web [145].

En el caso de los navegadores web que sus funcionalidades principales son la privacidad y la seguridad, podemos encontrar [146]:

- **Opera:** Ofrece una combinación de seguridad, privacidad y funcionalidad, pero no especifica el país al que se conecta la VPN.
- **TOR**: Es una opción segura para navegar de forma anónima, pero la protección avanzada reduce la velocidad.
- **Firefox:** Es un navegador rápido y seguro con opciones de privacidad personalizables, pero puede ser complejo de configurar.
- **Ungoogled Chromium:** Ofrece actualizaciones de seguridad frecuentes, pero se deben instalar manualmente.
- **Epic Privacy Browser:** Elimina los datos del usuario al cerrar el navegador, pero es difícil encontrar los archivos de código abierto que utiliza.
- **Brave:** Cuenta con un bloqueador de anuncios y rastreadores integrado, pero es un navegador relativamente nuevo.
- **Iridium:** Impide el seguimiento de los datos, pero no está disponible para dispositivos móviles.
- **Safari:** Utiliza el aislamiento para proteger las pestañas del malware, pero recopila datos para el mantenimiento del servicio.
- **Vivaldi:** Tiene una interfaz fácil de usar y permite personalizarla en gran medida, pero requiere registrarse con datos personales.
- **Chrome:** Mejora la seguridad del navegador con frecuencia, pero recopila y comparte datos con los anunciantes.

Igual que tenemos navegadores para ordenadores, también se han desarrollado navegadores web específicamente para teléfonos móviles.

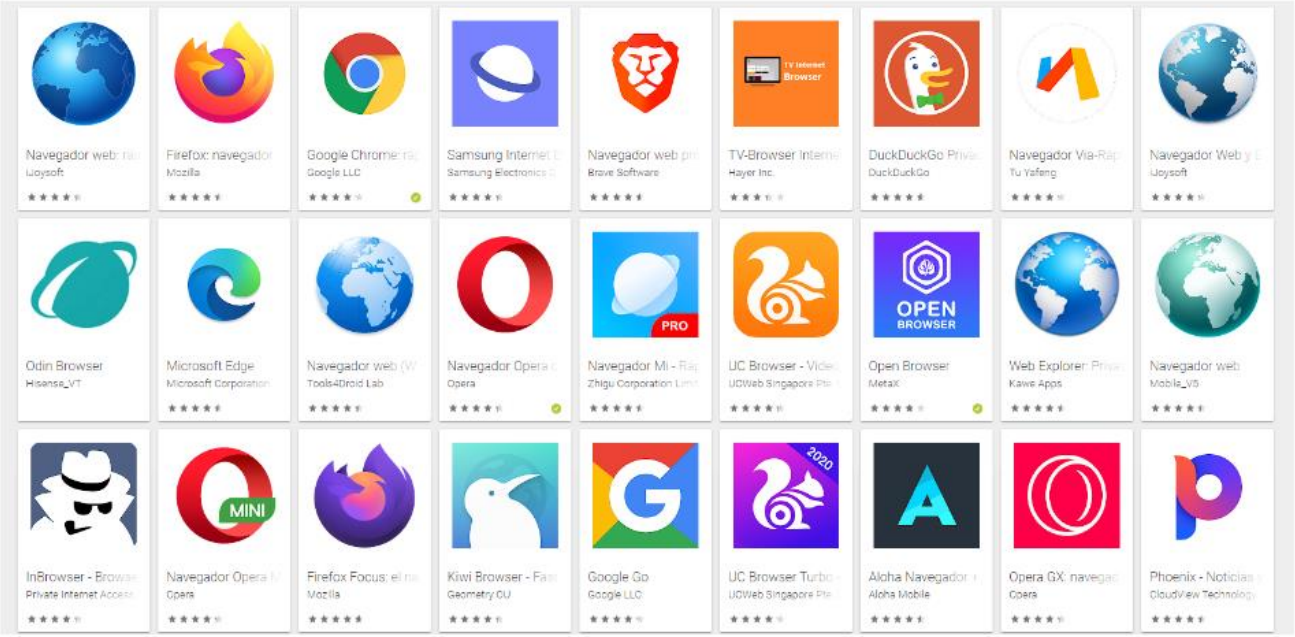


Figura 4-31 Muestra de navegadores web para móvil [147].

4.8 Seguridad y Privacidad en el DNS

El Sistema de Nombres de Dominio (DNS) es un protocolo esencial que conecta los nombres de los sitios web que visitamos con sus direcciones IP correspondientes. Es crucial prestar atención al servidor DNS que se utiliza, ya que cambiarlo puede tener un impacto considerable en la privacidad, la velocidad de navegación y la protección contra sitios web maliciosos [148].

Optar por un servidor DNS diferente al proporcionado por tu proveedor de servicios de Internet (ISP) puede brindarte ciertos beneficios en términos de privacidad. Al hacerlo, tienes la capacidad de ocultar tu historial de navegación, evitando que tu ISP acceda directamente a dicha información, ya que las solicitudes de traducción de nombres de dominio se realizan a través de un servidor diferente.

Además, el uso de un servidor DNS alternativo puede ayudar a eludir el bloqueo de sitios web específicos o la censura impuesta por tu proveedor o autoridades locales. Esto resulta especialmente útil en regiones con restricciones significativas en el acceso a Internet.

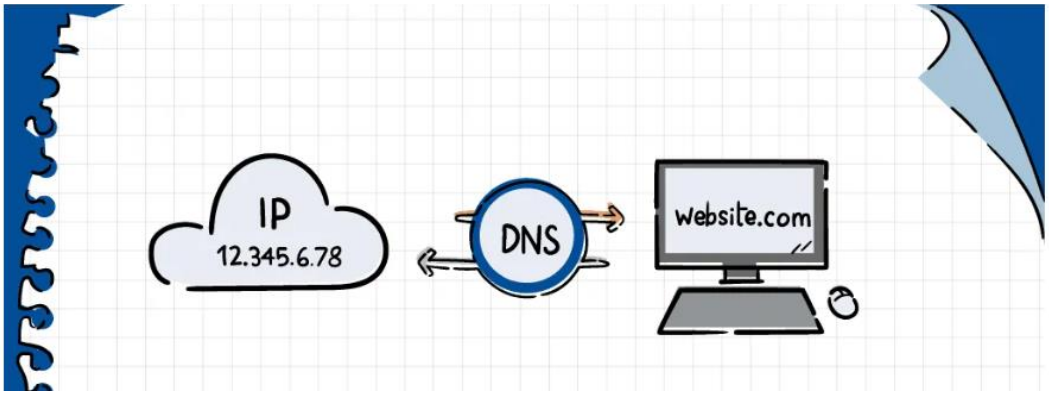


Figura 4-32 DNS (Domain Name Server) [149]

En algunos casos, los servidores DNS de terceros pueden ofrecer tiempos de respuesta más rápidos que los predeterminados proporcionados por tu ISP. Esta eficiencia se debe a su distribución en diversas ubicaciones geográficas y a su capacidad para gestionar un volumen considerable de solicitudes de manera más eficaz.

La selección cuidadosa de los servidores DNS puede mejorar de manera significativa tu privacidad y seguridad en línea. Es una medida importante para quienes buscan controlar su información personal, evitar restricciones de acceso y mejorar la velocidad de navegación.

Algunos ejemplos de servidores DNS de terceros populares son [150]:

- **Google Public** DNS: proporcionado por Google, ofrece servidores DNS rápidos y confiables. Sus direcciones IP son 8.8.8.8 y 8.8.4.4.
- **Cloudflare DNS:** ofrecido por Cloudflare, este servidor DNS destaca por su enfoque en la privacidad y la seguridad. Sus direcciones IP son 1.1.1.1 y 1.0.0.1.
- **OpenDNS:** proporcionado por Cisco, OpenDNS ofrece opciones gratuitas y de pago. Sus direcciones IP son 208.67.222.222 y 208.67.220.220.
- **Quad9:** un servicio DNS gratuito que se centra en la seguridad y bloquea sitios web maliciosos conocidos. Sus direcciones IP son 9.9.9.9 y 149.112.112.112.
- **DNS.Watch:** un servidor DNS independiente que promete privacidad y no registra información personal. Sus direcciones IP son 84.200.69.80 y 84.200.70.40.

Estos servidores DNS alternativos ofrecen diversas características, desde enfoques rápidos y confiables hasta un fuerte énfasis en la privacidad y la seguridad. La elección entre ellos puede depender de las preferencias individuales y las necesidades específicas de cada usuario.

Realizar pruebas y comparaciones es esencial para identificar el servidor DNS que proporciona la mejor velocidad para un caso de uso o una situación particular. Se pueden emplear herramientas como **DNS Benchmark** (https://www.grc.com/dns/benchmark.htm) para evaluar distintos servidores DNS y determinar cuál se adapta mejor a las necesidades requeridas.

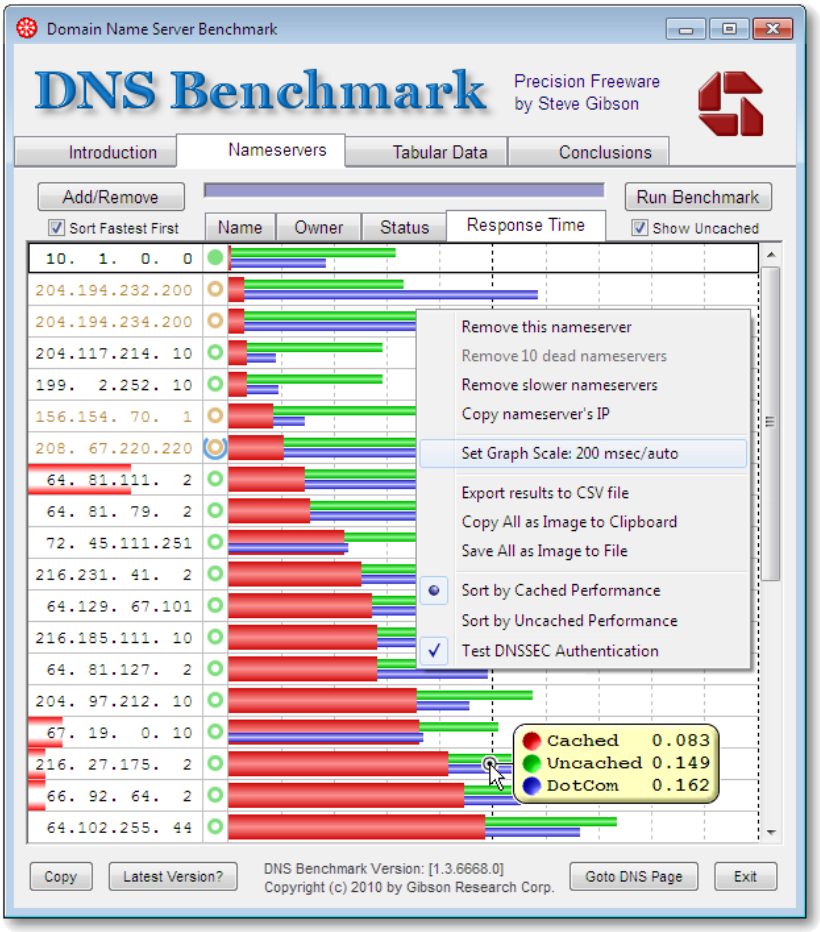


Figura 4-33 DNS Benchmark [151]

DNS over TLS / HTTPS / QUIC

Los términos "DoT," "DoH," y "DoQ" se refieren a tres protocolos distintos utilizados para mejorar la seguridad y privacidad en las comunicaciones DNS (Domain Name System) [152].

- **DNS over TLS (DoT):** DNS over TLS es un protocolo que cifra las consultas DNS utilizando el protocolo TLS (Transport Layer Security). Opera sobre el puerto 853 y proporciona una capa adicional de seguridad al ocultar las consultas DNS de posibles observadores, garantizando la privacidad y protegiendo contra la manipulación de datos.
- **DNS over HTTPS (DoH):** DNS over HTTPS es otro protocolo de seguridad para DNS que cifra las consultas utilizando el protocolo HTTPS. En lugar de utilizar el puerto DNS estándar (53), DoH utiliza el puerto HTTPS (443). Al implementar el cifrado HTTPS, DoH busca mejorar la privacidad y la seguridad, especialmente en entornos donde el tráfico DNS puede ser inspeccionado o manipulado.
- **DNS over QUIC (DoQ):** DNS over QUIC es un protocolo que utiliza QUIC (Quick UDP Internet Connections) para cifrar las consultas DNS. QUIC es un protocolo de transporte desarrollado por Google que se ejecuta sobre UDP (User Datagram Protocol) y está diseñado para mejorar la velocidad y la seguridad de las comunicaciones. DoQ busca proporcionar beneficios similares a DoT y DoH, pero utilizando QUIC como base.

Los tres protocolos comparten el objetivo común de mejorar la seguridad y la privacidad de las comunicaciones DNS mediante el cifrado de las consultas. La elección entre DoT, DoH o DoQ puede depender de las preferencias del usuario, la compatibilidad con la infraestructura existente y otros factores específicos del entorno de red.

Cloudflare proporciona una página web (https://www.cloudflare.com/es-es/ssl/encrypted-sni/?ref=ciberseguridad.blog#dns) que permite verificar el estado de la configuración. Al acceder a la página, al hacer clic en "Verificar mi navegador", se presentarán las medidas de seguridad implementadas y aquellas que aún no se han aplicado.

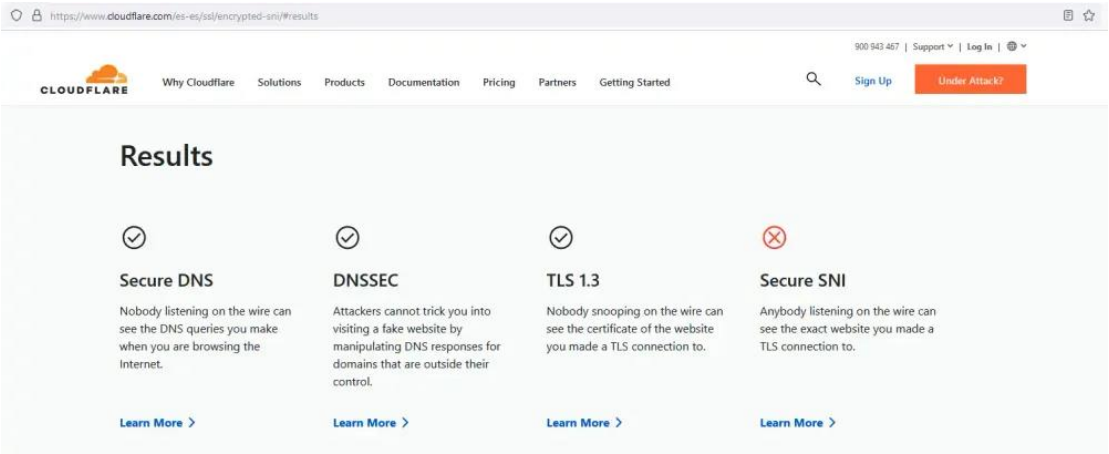


Figura 4-34 Comprobador DNS [153]

SNI y el estándar Encrypted Client Hello (ECH)

Toda la robustez de protocolos como SSL, TLS, QUIC, etc., tiene un punto vulnerable común conocido formalmente como **SNI (Server Name Indication**).

El **Server Name Indication (SNI)** es una extensión del protocolo de seguridad TLS que desempeña un papel crucial en la identificación de a qué nombre de host está tratando de conectarse el cliente antes de que se complete el proceso de handshaking. Básicamente, el SNI es una cabecera que

el cliente envía al servidor en texto plano al comienzo de la conexión, indicando el nombre de dominio.

Podemos pensar en el SNI como el equivalente de enviar un paquete a un edificio de apartamentos en lugar de a una casa. Cuando enviamos algo a una casa, la dirección de la calle es suficiente para que el paquete llegue a la persona correcta. Sin embargo, al enviar un paquete a un edificio de apartamentos, necesitamos el número de apartamento además de la dirección de la calle; de lo contrario, el paquete podría extraviarse o no ser entregado adecuadamente.

Muchos servidores web funcionan más como edificios de apartamentos que como casas, ya que alojan varios nombres de dominio. En esta situación, la dirección IP por sí sola no es suficiente para indicar a qué dominio específico está tratando de llegar un usuario. Esto podría resultar en la visualización incorrecta del certificado SSL por parte del servidor, lo que podría impedir o interrumpir una conexión segura mediante HTTPS.

Cuando varios sitios web comparten una única dirección IP en un servidor y cada uno tiene su propio certificado SSL, el servidor podría enfrentar dificultades para determinar qué certificado mostrar cuando un dispositivo cliente intenta conectarse de manera segura. Es aquí donde entra en juego la Server Name Indication (SNI). Esta extensión del protocolo TLS, utilizado en HTTPS, se incorpora al proceso de handshaking de TLS/SSL para garantizar que los dispositivos clientes puedan visualizar el certificado SSL correcto correspondiente al sitio web al que intentan acceder de forma segura [154]

Aunque los proveedores de servicios de Internet no pueden conocer las actividades específicas de sus clientes dentro de un sitio web ni acceder a sus datos, sí tienen la capacidad de identificar a qué sitios web acceden. Esto se evidencia, por ejemplo, en casos judiciales donde se bloquea el acceso a ciertos sitios web desde un país específico, limitando efectivamente el acceso a todos ellos mediante la identificación de las cabeceras SNI.

El **Encrypted Client Hello (ECH)** representa una mejora significativa en la privacidad del protocolo TLS, una parte fundamental de la infraestructura de Internet. ECH cifra todo el protocolo de enlace, salvaguardando así los metadatos privados y sensibles de la conexión TLS [155].

Como sucesor de ESNI, ECH oculta la Indicación del Nombre del Servidor (SNI), que se utiliza en la negociación de un handshake TLS. Esto implica que al visitar un sitio web en Cloudflare con ECH habilitado, únicamente el usuario, Cloudflare y el propietario del sitio pueden identificar qué sitio web se ha visitado [156].

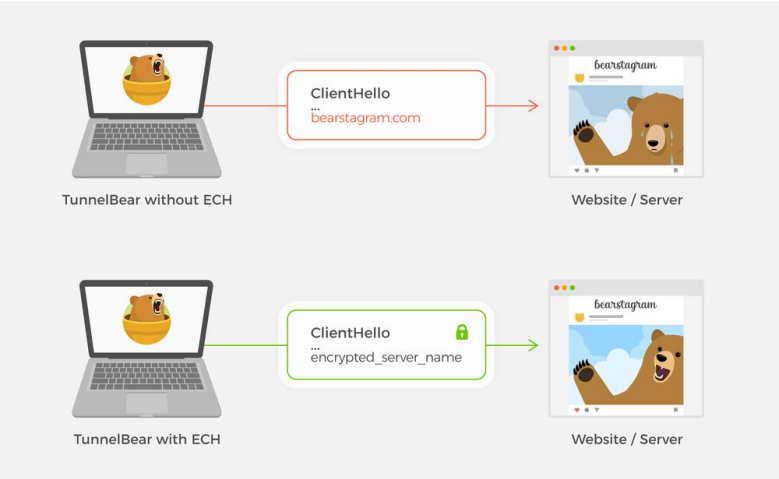


Figura 4-35 Funcionamiento ECH [156]

Además de fortalecer la privacidad, ECH sienta las bases para integrar futuras funciones de seguridad y mejoras de rendimiento en el protocolo TLS, al tiempo que minimiza el impacto en la privacidad de los usuarios finales.

Una ventaja adicional de ECH es su capacidad para preservar la conexión a nuestra infraestructura de backend, protegiéndola de terceros como proveedores de servicios de Internet, gobiernos o hackers que podrían intentar interceptar o observar la conexión de red. ECH se posiciona como la última tecnología disponible para mantener la seguridad y privacidad en estas comunicaciones [157].

4.9 Ejemplos de aplicaciones prácticas.

Como hemos comentado, navegar de forma anónima en Internet puede estar justificado por varias razones importantes relacionadas con la privacidad, la seguridad y la libertad. De tal forma que podamos proteger nuestra privacidad personal, para evitar la recopilación de hábitos de navegación, preferencias y datos personales, disminuimos el riesgo del robo de identidades reduciendo las posibilidades de ataques maliciosos. También es una manera de eludir la censura, garantizando la libertad de expresión y el acceso a la información y proporcionado una capa adicional de protección contra la vigilancia gubernamental.

Para ello debemos ser capaces de montar un sistema seguro que nos proporcione protección contra las amenazas mencionadas.

Alianza 5/9/14 ojos

Las expresiones "5 Eyes" (Cinco Ojos), "9 Eyes" (Nueve Ojos) y "14 Eyes" (Catorce Ojos) se refieren a alianzas de inteligencia entre países para compartir información y realizar vigilancia. Estas alianzas se formaron inicialmente para la cooperación en inteligencia de señales y monitoreo de comunicaciones. Cada número representa el número de países participantes en cada nivel de colaboración:

- **Cinco Ojos (5 Eyes):** Estados Unidos, Reino Unido, Canadá, Australia y Nueva Zelanda.
- **Nueve Ojos (9 Eyes):** Los mismos cinco países de Cinco Ojos, más Dinamarca, Francia, los Países Bajos y Noruega.
- **Catorce Ojos (14 Eyes):** Los nueve países de Nueve Ojos, más Alemania, Bélgica, Italia, España y Suecia.

Estas alianzas han sido motivo de preocupación en términos de privacidad y vigilancia, ya que implican el intercambio de información y colaboración en la monitorización de las comunicaciones a nivel internacional.

Estas alianzas de inteligencia estatales vigilan y comparten la actividad y datos de los usuarios de Internet para proteger la seguridad nacional. Esta cooperación comenzó con el Acuerdo UKUSA, un tratado para la cooperación conjunta en inteligencia de señales firmado al inicio de la Guerra Fría [158].

Los países participantes son conocidos por espiar a sus ciudadanos a través de diversos medios, obteniendo información sensible y privada que puede ser compartida con los otros miembros de la alianza [159].

Aparte de estas alianzas confirmadas, existe otro grupo de países que han sido sorprendidos o son sospechosos de intercambiar información con la Alianza de los Catorce Ojos. Estos incluyen a Israel, Japón, Singapur y Corea del Sur.

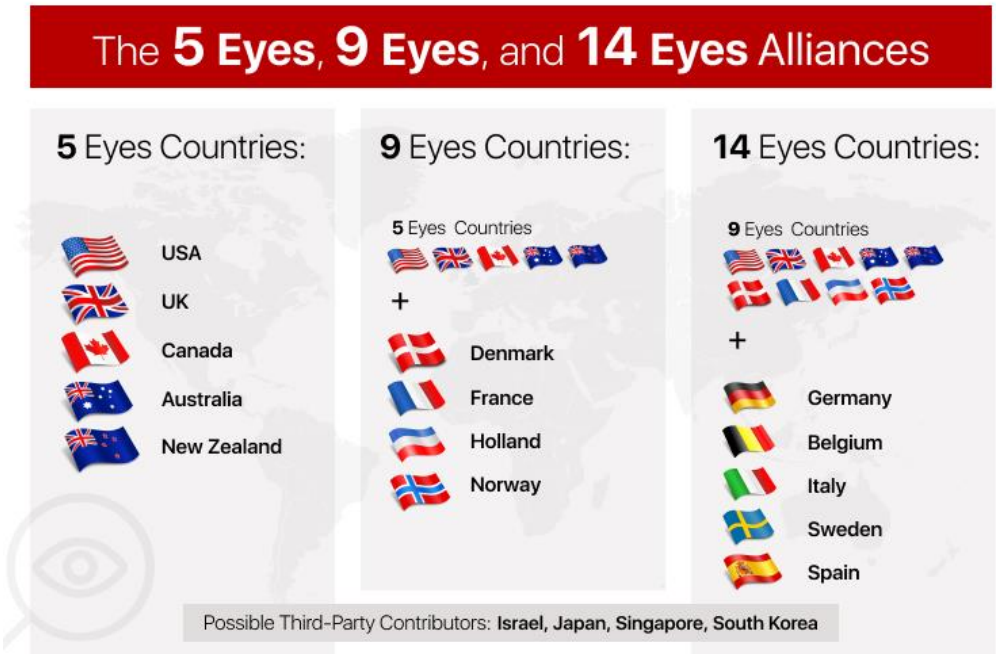


Figura 4-36 Países miembros de las Alianzas [159]

Una nueva expresión "Sexto Ojo" se utiliza para referirse a Japón como una posible adición a la Alianza de los Cinco Ojos. Aunque Japón no es oficialmente miembro de la alianza, mantiene estrechas relaciones de intercambio de inteligencia con los países de los Cinco Ojos.

En agosto de 2020, el Ministro de Defensa japonés expresó el deseo de una cooperación aún más estrecha con los Cinco Ojos y sugirió que Japón podría ser conocido como el "Sexto Ojo". Tanto Estados Unidos como el Reino Unido han mostrado interés en la posible participación de Japón. Sin embargo, la idea de que Japón sea el Sexto Ojo no es un reconocimiento oficial y sigue siendo en gran medida especulativa.

Estas alianzas tienen un gran número de sistemas de vigilancia masiva en funcionamiento, siendo algunos desconocidos para el público. No obstante, algunos programas han recibido una atención mediática significativa como ECHELON, PRISM, o XKeyscore [159].

4.9.1 VPN Segura

Dentro del sistema que deberíamos establecer, la primera medida es levantar una VPN segura. Una VPN debería seguir siendo la primera línea de defensa contra la vigilancia invasiva y la monitorización, pero la elección de un proveedor de VPN es sumamente importante. No todas las VPN están a salvo de las alianzas de los Ojos, algunas cooperan y seguirán haciéndolo con ellas.

Para establecer una VPN segura debemos tener en cuenta lo siguiente [158]:

- **Jurisdicción de VPN:** el país donde un proveedor de VPN está legalmente establecido. Dependiendo de cómo las autoridades relevantes supervisen el uso de VPN, muchos proveedores de VPN eligen establecer su empresa fuera de su país de residencia.
- **Ubicación del proveedor de VPN:** dónde se encuentra el proveedor de VPN como empresa, lo cual puede no ser lo mismo que dónde la empresa mantiene sus servidores VPN.
- **Ubicación del servidor de VPN:** dónde el proveedor de VPN ha decidido instalar servidores. Por lo general, un proveedor de VPN tiene servidores en múltiples ubicaciones, lo que te permite elegir entre muchas ubicaciones.

Por lo tanto, es muy recomendable no elegir un proveedor de VPN establecido en un país asociado con la Alianza de los Catorce Ojos para minimizar los problemas con la privacidad en línea. Además, las mejores VPN para la privacidad deben tener políticas estrictas de no almacenamiento de registros. Esto significa que no conservan ningún tipo de información identificativa sobre sus usuarios o su actividad en línea.

Algunos países (como China e Irán) poseen una regulación muy estricta respecto al uso de VPNs limitando su uso a una breve lista de VPNs "aprobadas por el gobierno", las cuales se puede asumir con seguridad que comparten datos con las autoridades. Otros han prohibido por completo el uso de VPNs.

Protocolos VPN

Existen varios protocolos utilizados en el ámbito de las redes privadas virtuales (VPN), cada uno con sus propias características, niveles de seguridad y aplicaciones específicas. Algunos de los protocolos de VPN más comunes son:

- **OpenVPN:** Es un protocolo de código abierto que es ampliamente considerado como uno de los protocolos VPN más seguros y confiables. Es compatible con la mayoría de los sistemas operativos y dispositivos, y utiliza cifrado de 256 bits para proteger los datos.
- **IPSec:** Es un conjunto de protocolos que se utilizan para cifrar y autenticar el tráfico de red. IPSec es compatible con la mayoría de los sistemas operativos y dispositivos, y es uno de los protocolos VPN más seguros disponibles.
- **L2TP:** Es un protocolo de túnel que se utiliza para conectar redes privadas virtuales. L2TP es compatible con la mayoría de los sistemas operativos y dispositivos, y utiliza cifrado de 256 bits para proteger los datos.
- **PPTP:** Es uno de los protocolos VPN más antiguos y menos seguros. Aunque es compatible con la mayoría de los sistemas operativos y dispositivos, no se recomienda su uso debido a las vulnerabilidades de seguridad conocidas.
- **SSTP:** Es un protocolo VPN propietario desarrollado por Microsoft. SSTP es compatible con Windows, Linux, macOS y dispositivos móviles, y utiliza cifrado de 256 bits para proteger los datos.
- **IKEv2:** Es un protocolo de túnel que se utiliza para conectar redes privadas virtuales. IKEv2 es compatible con la mayoría de los sistemas operativos y dispositivos, y es uno de los protocolos VPN más seguros disponibles.
- **WireGuard:** Es un protocolo VPN de código abierto que se ha vuelto cada vez más popular en los últimos años. WireGuard es compatible con la mayoría de los sistemas operativos y dispositivos, y utiliza cifrado de 256 bits para proteger los datos. Es conocido por su velocidad y eficiencia.

La elección del protocolo dependerá de las necesidades específicas, considerando factores como la seguridad, velocidad, compatibilidad y el entorno de uso (por ejemplo, conexiones móviles, acceso remoto, etc.). En general, se recomienda utilizar protocolos más seguros y modernos como OpenVPN o IKEv2/IPsec cuando la seguridad es una prioridad.

4.9.2 Sistema Operativo basado en la seguridad y privacidad

Los sistemas operativos que más seguridad y privacidad ofrecen son los sistemas operativos "live"[10] y ejecutarlo desde un USB o desde una máquina virtual, de esta manera, nuestra presencia sólo estará disponible mientras usemos este sistema operativo o máquina virtual.

Como ejemplo, mencionamos el sistema TAILS.

Figura 4-37 Logo de TAILS [160]

Tails es una buena solución porque ya incluye muchas de las funciones de seguridad necesarias para preservar el anonimato.

Además de la utilización de un sistema basado en la privacidad, se recomienda el "Full Disk Encryption" (FDE), que consiste en la completa encriptación del disco duro y de todos los archivos, el borrado seguro de los archivos por medio de aplicaciones específicas como Eraser, dban, fileshredder o priform.

Otra medida de seguridad adicional es la eliminación de los metadatos de los archivos y deshabilitar JavaScript.

Dentro de los metadatos de las fotos, están los datos Exif[11], que no se eliminan con el proceso de eliminación de metadatos usual. La siguiente dirección proporciona una versión online para ver los datos Exif de una imagen.

https://www.viewexifdata.com/?ref=ciberseguridad.blog

4.9.3 PGP para intercambio de archivos

PGP (Pretty Good Privacy) se ha convertido en un fundamento de la privacidad y seguridad en Internet porque posibilita el envío de mensajes cifrados a alguien sin necesidad de compartir el código de antemano [161].

PGP emplea una amalgama de técnicas de cifrado, como el hash, la compresión de datos, la criptografía simétrica de clave privada y la criptografía asimétrica de clave pública, con el fin de garantizar la seguridad de los datos. Actualmente, las versiones más seguras de PGP utilizan claves de longitud de 4096 bits [162].

El cifrado PGP se puede aplicar para proteger archivos de texto, correos electrónicos, archivos de datos, directorios y particiones de disco.

Las versiones más recientes de PGP son esencialmente impenetrables, siempre y cuando se utilicen de manera adecuada. Aunque hay algunas vulnerabilidades teóricas en las versiones anteriores, no se tiene conocimiento público de que las versiones actuales puedan ser comprometidas mediante tecnologías contemporáneas y las últimas técnicas de criptoanálisis [161].

---

[10] Sistema operativo que puede ejecutarse sobre otro sistema operativo.

[11] Exchangeable image file format (Formato de archivo de imagen intercambiable) es una especificación para formatos de archivos de imagen usado por las cámaras digitales.

4.9.4 Esteganografía para el intercambio de información

La esteganografía, una técnica milenaria que se ha adaptado a la era digital, ofrece una fascinante capacidad para ocultar información dentro de otros datos, ya sean imágenes o archivos de audio, con el objetivo de eludir la detección. Su presencia en la actualidad se extiende entre diversos actores, desde espías hasta piratas informáticos malintencionados, pasando por activistas de derechos humanos y disidentes políticos. De hecho, la esteganografía digital ha emergido como uno de los componentes esenciales en las cajas de herramientas de estos grupos, proporcionando una capa adicional de sigilo para la transmisión de mensajes y datos sensibles [163].

Figura 4-38 Esteganografia sobre imagen [164]

A pesar de su antigüedad, la esteganografía sigue evolucionando en consonancia con los avances tecnológicos. Su raíz histórica se remonta al siglo III a.C., cuando Filón de Bizancio ya escribía con tintas invisibles, pero su aplicación contemporánea se ha vuelto cada vez más sofisticada. La esteganografía, en esencia, busca ocultar información, imágenes o mensajes de audio dentro de otra pieza de información, imagen o audio, utilizando técnicas diversas y avanzadas.

En el mundo digital, existen varias formas de esteganografía, cada una adaptada a distintos propósitos. Entre estas variantes se encuentran la esteganografía pura, la esteganografía de clave secreta y la esteganografía de clave pública. Lo distintivo de la esteganografía es que difiere de la criptografía en su enfoque; mientras que la criptografía implica codificar datos utilizando claves, la esteganografía se centra en ocultar datos de manera inteligente, sin necesidad de codificación o claves.

Las técnicas utilizadas en la esteganografía digital son diversas y abarcan diferentes tipos de archivos, desde documentos hasta imágenes, vídeos y archivos de audio. Entre estas técnicas se incluyen el enmascaramiento, algoritmos de compresión de datos y métodos de sustitución, que permiten integrar información de forma casi imperceptible en el medio de transporte.

Es crucial destacar que, a pesar de su utilidad para ocultar información de miradas indiscretas, la esteganografía no se concibe como una medida segura para proteger la información. Su función principal radica en la ocultación, no en la protección. Si bien puede ser eficaz para eludir a posibles fisgones, no debe considerarse como una forma infalible de resguardar la información. En este sentido, la implementación de la esteganografía debe ir acompañada de otras medidas de seguridad más robustas para garantizar la protección integral de la información sensible.

4.9.5 Navegador TOR y VPN

La combinación de una VPN y Tor es la mejor manera de proteger la privacidad en línea. Una VPN cifra el tráfico, lo que dificulta que los rastreadores y los gobiernos rastreen. TOR oculta la dirección IP, lo que hace que sea aún más difícil que la identificación.

La VPN, con su protocolo de cifrado, impedirá que nodos maliciosos vean la dirección IP y actividad, al tiempo que oculta el uso de Tor ante el proveedor de servicios de Internet (ISP) y

organismos de vigilancia. Evitar alertar sobre la actividad en Internet es crucial. Además, esta configuración permitirá acceder a sitios web que bloquean usuarios de TOR [165].

TOR a través de la VPN

El uso de TOR a través de la VPN conecta primero con la VPN antes de iniciar TOR. De esta manera, la VPN cifra el tráfico antes de que atraviese la red de TOR, manteniendo oculto el uso de TOR ante el ISP.

Al usar Tor a través de la VPN, el proveedor de VPN no puede ver los datos que se envían a través de TOR, pero puede detectar la conexión a dicha red. El nodo de entrada de TOR no tiene acceso a la IP real; en cambio, visualiza la IP del servidor VPN, mejorando el anonimato.

Sin embargo, al salir de la red de Tor, el tráfico no está cifrado. TOR a través de la VPN no protege contra nodos de salida maliciosos, por lo se debe ser cauteloso al enviar información delicada a través de la conexión.

Los casos de uso pueden ser:

- Ocultar el uso de TOR al ISP y a organismos de vigilancia.
- Mantener oculto el tráfico al proveedor de VPN.
- No enviar información personal, como credenciales de inicio de sesión, a través de la conexión.

VPN a través de TOR

La metodología de conectar la VPN a través de TOR opera en sentido opuesto a TOR a través de la VPN. En este caso, primero se conecta a Internet y luego se inicia sesión en la VPN desde la red de TOR. Este enfoque es un tanto más técnico y complejo, ya que requiere la configuración específica del cliente VPN para que funcione con TOR.

En lugar de acceder directamente a Internet, el nodo de salida de TOR dirige el tráfico hacia el servidor VPN. Esto elimina el riesgo de nodos de salida maliciosos, dado que el tráfico se desencripta después de abandonar la red TOR.

Aunque el nodo de entrada de Tor aún puede visualizar la IP real, la VPN únicamente verá la dirección del nodo de salida. El proveedor de servicios de Internet (ISP) no podrá detectar la conexión a una VPN, pero sí notará que el uso de TOR. Gracias a la capacidad de seleccionar el servidor remoto de la VPN, también se facilita eludir la censura y las restricciones geográficas con este método.

Los casos de uso de esta solución son:

- Resguardar la conexión contra nodos de salida malintencionados.
- Que el ISP no detecte el uso de una VPN.
- Transmisión de información delicada a través de la conexión, como credenciales de inicio de sesión y mensajes privados.
- Superar bloqueos geográficos.

TOR portable

"TOR Portable" hace referencia a una versión portátil del navegador Tor. La versión "portable" generalmente significa que el navegador está configurado para ser ejecutado desde una unidad USB o cualquier otro dispositivo de almacenamiento portátil, sin necesidad de instalación en el sistema operativo principal. Esto permite llevar el navegador Tor contigo y usarlo en diferentes computadoras sin dejar rastros en el sistema host.

4.9.6 Nodos Puente de TOR

La Red TOR consta de dos tipos de nodos de retransmisión: nodos de retransmisión normales y nodos de retransmisión puente. Los nodos de retransmisión normales están enumerados en el directorio principal de TOR, y las conexiones a ellos pueden ser fácilmente identificadas y bloqueadas por censores [166].

Los nodos puente de TOR son relés secretos de la red TOR y son una parte importante de la infraestructura diseñada para ayudar a los usuarios a eludir la censura y la vigilancia en línea ya que mantienen la conexión oculta. Se emplean como el primer eslabón en la cadena de TOR cuando la conexión a esta red está bloqueada o cuando el uso de TOR podría levantar sospechas por parte de alguien que esté monitorizando una conexión a Internet.

La mayoría de los transportes conectables, como obfs4, dependen del uso de estos nodos "puente". Estos nodos, son operados por voluntarios al igual que los nodos de retransmisión normales de TOR. Los nodos puente no están públicamente listados en el directorio principal de TOR, como los nodos de retransmisión normales. Su función principal es actuar como puntos de acceso ocultos que no son fácilmente identificables por los censores.

Cuando un usuario intenta conectarse a la red Tor desde una ubicación donde la censura es estricta, los nodos puente pueden ser utilizados para establecer la conexión de forma más discreta. Esto ayuda a evitar la detección y bloqueo por parte de la censura, ya que los nodos puente no son predecibles y su información no se encuentra en el directorio principal.

Los nodos puente en la red TOR son una medida para mejorar la resistencia contra la censura y permitir que los usuarios accedan a la red de forma más segura y privada, evitando bloqueos y restricciones impuestas por terceros.

4.9.7 Nodos Puente Ofuscados de TOR

Los nodos ofuscados (*Pluggable Transport*) son extensiones opcionales del protocolo TOR que proporcionan mecanismos adicionales de ofuscación diseñados para hacer que el tráfico de Tor sea menos detectable y, por lo tanto, menos susceptible a la censura. Estos Transports modifican la apariencia del tráfico de Tor para que parezca algo diferente, como tráfico HTTPS o tráfico no cifrado. La idea es dificultar la identificación del tráfico de Tor, lo que hace que sea más difícil para los sistemas de censura bloquearlo [167].

Esta solución se ha desarrollado debido a que un análisis de tráfico, incluso utilizando los puentes de TOR, usando un DPI (inspección profunda de paquetes) puede identificar los diferentes tráficos de Internet por protocolos, de esta manera indicando el uso del protocolo TOR [160].

Actualmente, el proyecto TOR soporta algunos *pluggable transports*.

- **Obfsproxy (Protocolo de Ofuscación):** Obfsproxy es un marco que permite la implementación de varios protocolos de ofuscación. Estos protocolos están diseñados para hacer que el tráfico de Tor parezca algo diferente, como tráfico HTTPS o tráfico de chat cifrado.
- **Meek:** Meek utiliza servicios en la nube como intermediarios para enrutar el tráfico de Tor. El tráfico se camufla como tráfico HTTPS hacia servicios como Amazon CloudFront o Google App Engine, lo que hace que sea difícil distinguirlo del tráfico web regular.
- **Snowflake:** Snowflake aprovecha la ayuda de voluntarios que ejecutan navegadores web con un proxy especial. Los usuarios de Tor pueden conectarse a través de estos proxies, y el tráfico se camufla como tráfico web normal.

- **ScrambleSuit:** ScrambleSuit es otro protocolo de ofuscación que se utiliza como un pluggable transport en Tor. Su objetivo es dificultar la detección del tráfico de Tor haciéndolo parecer como tráfico no cifrado.

4.9.8 Anonimato mediante el uso de proxies

La mayoría de los proveedores de listas de proxies y aplicaciones de software categorizan los servidores proxy en tres categorías según el nivel de anonimato que proporciona cada servidor proxy. Por lo general, estas categorías van desde el Nivel 1 hasta el Nivel 3. Los proxies marcados como Nivel 1 indican el nivel máximo de anonimato que un proxy podría tener, mientras que los Niveles 3 indican el mínimo. Algunas listas de proxies etiquetan sus categorías de manera diferente; por ejemplo, los proxies de Nivel 1 suelen llamarse Elite proxies o incluso High Anonymous proxies. Los proxies de Nivel 2 se suelen denominar Anonymous Proxies, y los proxies de Nivel 3 se conocen como Transparent Proxies [168].

La única manera de ocultar la identidad al servidor de destino es enviar la solicitud a través de un intermediario, un intermediario que realizará la solicitud en el nombre del cliente y se la devolverá. Este es el concepto básico de un servidor proxy. Cuando un cliente utiliza un proxy, todas sus solicitudes se envían a un servidor proxy (intermediario) que luego obtiene el recurso solicitado en nombre del cliente y se lo devuelve. Si ese servidor proxy reenvía información adicional a un servidor de destino, como identificarse como un servidor proxy o revelar la dirección IP del cliente, eso determina el nivel de anonimato que tiene ese proxy.

Los servidores proxy pueden tener diferentes niveles de anonimato, y estos niveles definen cuánta información sobre el cliente (usuario) se revela al servidor de destino. Los tres niveles comunes de anonimato de proxy son:

- **Proxy Transparente (Nivel 3):** Un proxy transparente, también conocido como proxy interceptador o forzado, no modifica los encabezados de la solicitud del cliente. Pasa la dirección IP original del cliente al servidor de destino, haciendo que la presencia del cliente sea detectable. Los proxies transparentes a menudo se utilizan con fines de almacenamiento en caché y filtrado de contenido. Los proxies transparentes no son completamente inútiles para todos los propósitos, porque técnicamente la dirección IP enviada está "oculta" al servidor de destino, solo que sería muy fácil de descubrir.
- **Proxy Anónimo (Nivel 2):** Un proxy anónimo a veces conocido como proxie distorsionante, oculta la dirección IP del cliente al servidor de destino. Aunque no pasa la dirección IP del cliente en los encabezados, aún puede revelar que se está utilizando un proxy. El nivel de anonimato puede variar entre diferentes proxies anónimos, y algunos pueden agregar encabezados que indican el uso de un proxy.
- **Proxy Elite o de Alto Anonimato (Nivel 1):** Los proxies elite, también conocidos como proxies de alto anonimato, ofrecen el nivel más alto de anonimato. No solo ocultan la dirección IP del cliente, sino que tampoco agregan ningún encabezado que indique el uso de un proxy. Los proxies elite dificultan que los sitios web detecten que se está utilizando un proxy, ofreciendo un mayor nivel de privacidad.

La elección del nivel de anonimato del proxy depende de las necesidades específicas del usuario y del nivel de privacidad que requiera. Los proxies transparentes suelen ser menos adecuados para fines de privacidad, mientras que los proxies anónimos y elite ofrecen grados variables de anonimato.

Independientemente del nivel de anonimato del proxy, los usuarios deben considerar la confiabilidad del proveedor de proxy. Si el servidor proxy está comprometido o es malicioso, podría comprometer la privacidad del usuario. Además, el uso de cifrado, como HTTPS, en combinación con un proxy puede mejorar la seguridad y la privacidad de la conexión.

5 ELIMINACIÓN DE HUELLA DIGITAL

5.1 Cómo se genera la huella digital

La huella digital en Internet es un conjunto de información y datos generados cuando un usuario navega por la red y que les permite identificarlos de manera única. Gracias a diferentes tecnologías, esta información se puede recopilar de varias formas [169]:

- A través del uso de cookies.
- Mediante la dirección IP del dispositivo que se utiliza para navegar.
- A través de la información que el usuario proporciona voluntariamente al interactuar con diferentes sitios web.
- Mediante otros datos recopilados automáticamente mientras el usuario está navegando por Internet.

Esta información se puede utilizar para seguir la actividad en línea del usuario y proporcionar contenido o anuncios personalizados.

5.1.1 Tipos de cookies

Las cookies son archivos pequeños que contienen datos y se almacenan en el ordenador del usuario cuando visita una página web. Estos archivos guardan cierta información sobre el usuario, como su comportamiento de navegación en Internet o las credenciales de usuario [170].

Analistas web y desarrolladores utilizan cookies para obtener información sobre los usuarios y mejorar su experiencia en un sitio web.

Una de las funciones fundamentales de las cookies web es recordar los accesos y las credenciales de los usuarios en las páginas web. Esencialmente, estas cookies permiten que la página web identifique y recuerde al usuario, evitando que tenga que ingresar sus credenciales cada vez que accede al sitio web.

La Agencia General de Protección de Datos, establece las siguientes categorías de cookies que se pueden utilizar [171]:

- **Cookies Propias o de Terceros:** las cookies propias son gestionadas desde el dominio del webmaster, el titular legal de las mismas. En cambio, las cookies de terceros son administradas por entidades externas al editor, y se utilizan con fines analíticos u otros propósitos específicos.

- **Cookies Persistentes y de Sesión:** las cookies de sesión almacenan información que se utiliza solo durante la visita del usuario al sitio web. Por otro lado, las cookies persistentes se guardan de forma continua y se accede a ellas cuando sea necesario.
- **Cookies Técnicas, de Personalización, de Análisis y Publicitarias:** las cookies técnicas supervisan el tráfico y recopilan datos de comunicación. Las cookies de personalización crean una experiencia personalizada en la página, adaptada al idioma del usuario, por ejemplo. Las cookies de análisis registran datos para comprender el comportamiento del usuario y crear perfiles de visitantes. Por último, las cookies publicitarias se utilizan para gestionar espacios de promoción de productos o servicios.

Desde un punto de vista técnico podríamos identificar cuatro tipos generales de cookies y unas 13 subclases [171].

Tipos de Cookies según el Consentimiento

Una primera clasificación de los diferentes tipos de cookies en un sitio web se basa en el consentimiento y sirve para agrupar el resto de las cookies en Internet. Nos referimos a las cookies exceptuadas y las cookies no exceptuadas.

- **Cookies Exceptuadas:** las cookies exceptuadas son aquellas que no requieren el consentimiento previo del usuario para ser utilizadas. Estas incluyen las cookies técnicas y las cookies de personalización (de las cuales hablaremos más adelante). Estas cookies se consideran necesarias para el correcto funcionamiento del sitio web y, por lo tanto, pueden utilizarse sin el consentimiento de los usuarios.
- **Cookies No Exceptuadas:** por otro lado, el resto de las cookies son las cookies no exceptuadas y requieren el consentimiento expreso y previo de los usuarios para ser utilizadas. Esto significa que el aviso y la política de cookies deben informar sobre ellas y, además, permitir su aceptación o rechazo de forma sencilla.

  Ejemplos de tipos de cookies no exceptuadas incluyen las cookies de terceros, las cookies de seguimiento, las cookies publicitarias, entre otras.

Tipos de Cookies según la Entidad que las Gestiona

En función de la entidad que administra las cookies, es decir, el dominio o servidor al que las cookies son enviadas y donde se procesan los datos que recopilan, podemos distinguir entre:

- **Cookies Propias:** Las cookies propias, también conocidas como cookies de origen, son establecidas por el sitio web que el usuario visita directamente. Los datos recopilados mediante estas cookies se utilizan para propósitos como calcular las páginas vistas, las sesiones y el número de usuarios.

  Por lo general, los editores tienen acceso a los datos recopilados mediante estas cookies, que luego pueden compartirse con anunciantes o agencias para orientar la publicidad. Además, herramientas de análisis como Google Analytics utilizan cookies propias para comprender el comportamiento del usuario y presentar informes detallados en forma tabular o gráfica para que los editores los comprendan mejor.
- **Cookies de Terceros:** Las cookies de terceros son establecidas por dominios que los usuarios no visitan directamente. Esto ocurre cuando los editores incorporan elementos de terceros, como chatbots, complementos sociales o anuncios, en sus sitios web.

  Una vez instaladas, estas cookies también rastrean a los usuarios y almacenan su información para orientar anuncios y publicidad conductual. Por ejemplo, si se agrega un enlace de YouTube a un blog y un usuario hace clic en él, se añadirá una cookie de

YouTube al navegador del usuario. Esta cookie puede rastrear al usuario hasta que caduque.

Es importante destacar que las cookies de terceros podrían estar en proceso de desaparición, ya que muchos navegadores las bloquean por defecto en la actualidad.

Tipos de Cookies según el Plazo de Tiempo que Permanecen Activas

También podemos clasificar los tipos de cookies según el tiempo que permanecen activas en el navegador del usuario:

- **Cookies de Sesión:** las cookies de sesión caducan inmediatamente o pocos segundos después de que el usuario abandone el navegador web. Estas cookies se utilizan en sitios web de comercio electrónico para recordar los productos colocados en el carrito por el usuario, mantener a los usuarios conectados y calcular cada sesión de usuario con fines analíticos.

  Por ejemplo, si un sitio web de comercio electrónico no utiliza cookies de sesión, los artículos agregados al carrito se eliminarán cuando el usuario llegue a la página de pago. Además, el servidor olvidará al usuario y lo tratará como un visitante completamente nuevo.

- **Cookies Persistentes:** como su nombre indica, las cookies persistentes permanecen en el navegador del usuario durante un período prolongado. Por lo general, estas cookies tienen una fecha de vencimiento que puede oscilar entre un segundo y diez años.

  Los editores utilizan cookies persistentes para rastrear a un usuario individual y su interacción con el sitio web.

  Para verificar si tu navegador tiene cookies persistentes, puedes hacer lo siguiente: si has iniciado sesión en Gmail en el navegador, cierra la pestaña y reinicia tu dispositivo. Luego, vuelve a abrir el mismo navegador y visita el mismo servicio o cuenta (Gmail). Si aún estás conectado, significa que tienes cookies persistentes almacenadas en el navegador, las cuales fueron preservadas por el servicio de correo de Google.

Tipos de Cookies según su Finalidad

Las cookies también pueden clasificarse según la finalidad para la que recopilan los datos de los usuarios, distinguiendo entre:

- **Cookies Técnicas:** las cookies técnicas permiten al usuario navegar por una página web, sitio en línea o aplicación y utilizar sus servicios o funciones. Por ejemplo, proporcionan acceso a zonas restringidas, recuerdan los elementos de un carrito, facilitan el proceso de inscripción y almacenan contenidos para la difusión de vídeos, entre otros.

- **Cookies de Personalización:** estas cookies almacenan las preferencias de experiencia del usuario, como el idioma seleccionado.

- **Cookies de Análisis:** las cookies analíticas se utilizan para evaluar el uso de un sitio web y pueden rastrear a un usuario individual, pero solo en la medida en que les permita navegar por el sitio. No se utilizan para segmentar anuncios, a diferencia de otro tipo de cookies.

- **Cookies Publicitarias:** estas cookies están diseñadas específicamente para recopilar información sobre el usuario en su dispositivo para enviar publicidad basada en temas relevantes que le interesan. Los anunciantes colocan estas cookies en un sitio web con el permiso del operador del sitio. La información recopilada puede compartirse con otros anunciantes para mejorar el rendimiento de sus anuncios. Además, se utilizan para acumular estadísticas sobre el rendimiento de los anuncios en diversos sitios web. Por lo general, son cookies persistentes de terceros que siguen al usuario mientras navega por

otros sitios web. Ejemplos incluyen cookies de redes sociales que rastrean a los usuarios en la web para mostrarles anuncios en plataformas sociales.

- **Cookies de Publicidad Conductual:** la publicidad conductual en línea implica recopilar información sobre la actividad en línea del usuario para mostrar anuncios o contenido que se considera relevante. Las empresas utilizan cookies para inferir los intereses del usuario según las páginas visitadas, el contenido seleccionado y otras acciones en línea.
- **Cookies de Complementos:** las cookies de complementos se generan al usar complementos de proveedores externos de contenido. Permiten a los usuarios acceder a contenidos o servicios proporcionados por terceros, como ubicaciones en Google Maps, conexiones con redes sociales o acceso a YouTube.
- **Cookies del Reproductor Multimedia:** las cookies del reproductor de contenido facilitan la reproducción de audio o video en línea. Por ejemplo, permiten la reproducción automática de un video cuando un usuario navega por un sitio web.

Ejemplos de cookies de Google:

Algunas de las cookies que genera Google son:

- **NID:** esta cookie de funcionalidad es utilizada por Google para crear un ID único que permite recordar nuestras preferencias y otra información personalizada.
- **pm_sess:** esta cookie de seguridad se emplea para prevenir el spam y garantizar la integridad del sitio.
- **_ga:** se trata de la principal cookie utilizada por Google Analytics para distinguir a los usuarios en los sitios web.
- **IDE y ANID:** son dos cookies publicitarias de Google que se utilizan para mostrar sus anuncios en sitios de terceros. Estas cookies son empleadas en estrategias publicitarias en línea.

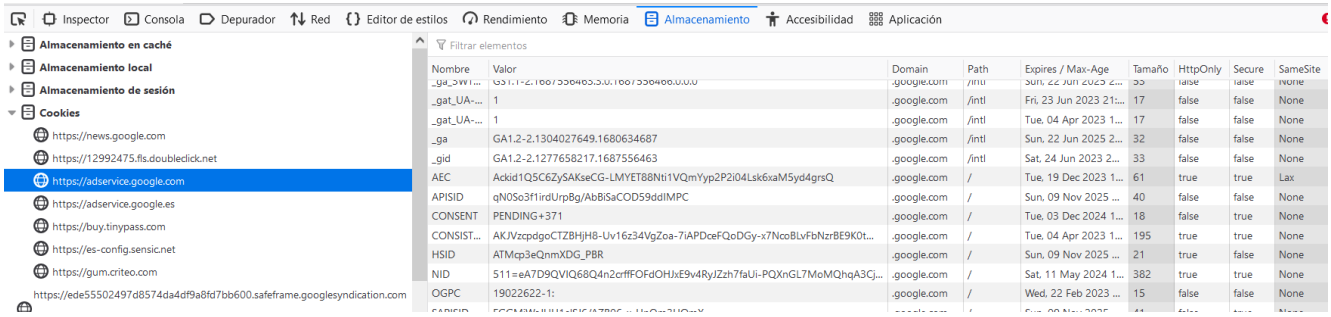


Figura 5-1 Ejemplo de cookies de Google

5.1.2 Supercookies

Las cookies son archivos que muchas páginas web usan para identificar a los usuarios y guardar información sobre sus hábitos y preferencias de navegación. También sirven para mantener abierta una sesión o para mostrar anuncios personalizados según lo que se busque en Internet. También son las culpables de que cada vez que se busca un producto se vean docenas de anuncios sobre el mismo allá en donde se navegue, lo que se conoce como *'retargeting'*.

Pero hay un tipo de cookie que no se puede borrar fácilmente desde las opciones del navegador. Se llaman supercookies, permacookies, zombie cookies, etc, y son una forma de rastrear la actividad online sin que uno se dé cuenta. Incluso si se usa el modo privado, estas cookies pueden seguir al usuario y recopilar datos [172].

Las supercookies son similares a las cookies de seguimiento en el sentido de que permiten a un rastreador compilar las visitas a diferentes sitios web. La diferencia clave es que, a diferencia de las cookies, el navegador nunca fue diseñado para almacenar supercookies. En cambio, las compañías de seguimiento han encontrado formas de abusar de otras características del navegador no relacionadas para colocar secretamente sus supercookies. Esto a menudo hace que sea más difícil para el navegador eliminar o bloquear las supercookies que bloquear las cookies normales de seguimiento [173].

Supercookie de TrustPid

TrustPid, una plataforma de publicidad digital creada por cuatro de las mayores operadoras de Internet de Europa (Telefónica, Orange, Vodafone y Deutsche Telekom), tiene como objetivo apoyar las actividades de marketing y publicidad digital de marcas y editores [174].

TrustPid funciona mediante un token que se genera a nivel del proveedor de servicios de Internet móvil (ISP) y que asigna una IP fija a cada usuario. Con este token, se pueden crear perfiles anonimizados para anunciantes y editores, que pueden ofrecer publicidad más personalizada y relevante.

Sin embargo, TrustPid también plantea algunas cuestiones sobre la privacidad y el consentimiento de los usuarios. A diferencia de las cookies, que se pueden borrar o bloquear desde el navegador, TrustPid sólo se puede desactivar desde su portal de privacidad. Además, el servicio se renueva automáticamente cada 90 días, por lo que hay que estar atento para volver a desactivarlo si no se quiere ser rastreado.

Cookies zombie

Las cookies zombie son un tipo de super cookie que no se almacenan en el navegador, sino en otras partes del mismo, como el historial de navegación, la caché o los códigos de color RGB que se usan en el navegador. Se instalan de tal forma que, aunque las eliminemos del almacén de cookies, siguen estando en el equipo.

Las cookies zombie no solo se pueden colocar en un solo lugar del almacenamiento local, sino en varios distintos. Si no borramos todas ellas, pueden volver a la vida una vez más.

Estas cookies son muy difíciles de eliminar, ya que no se pueden borrar con las opciones habituales del navegador.

La "cookie zombie," llamada así por su inmortalidad y su capacidad para "resucitar" cada vez que regresamos al sitio que la instaló (término acuñado por el abogado Joseph H. Malley), permanece almacenada y oculta en algún lugar del almacenamiento local de la computadora. Así, cuando el sitio de origen la detecta, la reactiva junto con cualquier otra cookie relacionada, reiniciando el proceso de recopilación de datos y rastreo de nuestras actividades [175].

Cookies Flash

Las cookies Flash son archivos exclusivos de animaciones Flash que almacenan temporalmente información sobre texto, imágenes y/o animaciones utilizadas por el complemento.

La complicación con esas cookies específicas de Flash radica en que se están empleando con objetivos distintos: las agencias publicitarias las utilizan para intentar recopilar información sobre nuestros hábitos de navegación. El navegador no puede gestionar directamente estas cookies, ya que están completamente controladas por un complemento (add-on). Por ende, es responsabilidad del usuario supervisar y eliminar de manera periódica la información almacenada en las cookies Flash [176].

Evercookies

Un evercookie es un tipo de supercookie que se crea usando una API de Javascript que identifica y reproduce las cookies que se han borrado intencionalmente en el almacenamiento del navegador del

cliente. Fue creado por Samy Kamkar en 2010 para demostrar la posible infiltración de los sitios web que usan el respawning.

Un evercookie almacena los datos de la cookie en varios mecanismos de almacenamiento que están disponibles en el navegador, como el almacenamiento local, la caché, el historial, los favoritos, los datos de Flash, los datos de Silverlight, etc. Si se borra alguno de estos mecanismos, el evercookie los sincroniza y los regenera. De esta forma, el evercookie es muy difícil de eliminar y puede identificar al cliente incluso después de que haya eliminado las cookies normales, las cookies de Flash (Local Shared Objects o LSOs) y otras [177].

Dentro de las categorías de supercookies, podemos además de las estudiadas, encontrar otra gran variedad de las mismas como son las Canvas Fingerprinting, ETag Tracking, Silverlight Isolated Storage, HSTS Supercookies, etc.

5.1.3 Información que proporciona una dirección IP

Una dirección IP (Protocolo de Internet) es un número único asignado a cada dispositivo conectado a una red que utiliza el protocolo IP para la comunicación. La información que se puede derivar de una dirección IP incluye:

- **Ubicación geográfica general:** la dirección IP puede proporcionar información sobre la ubicación geográfica aproximada de un dispositivo. Sin embargo, esta información no es siempre precisa y solo puede identificar la ubicación de la red, no el lugar físico exacto del dispositivo.

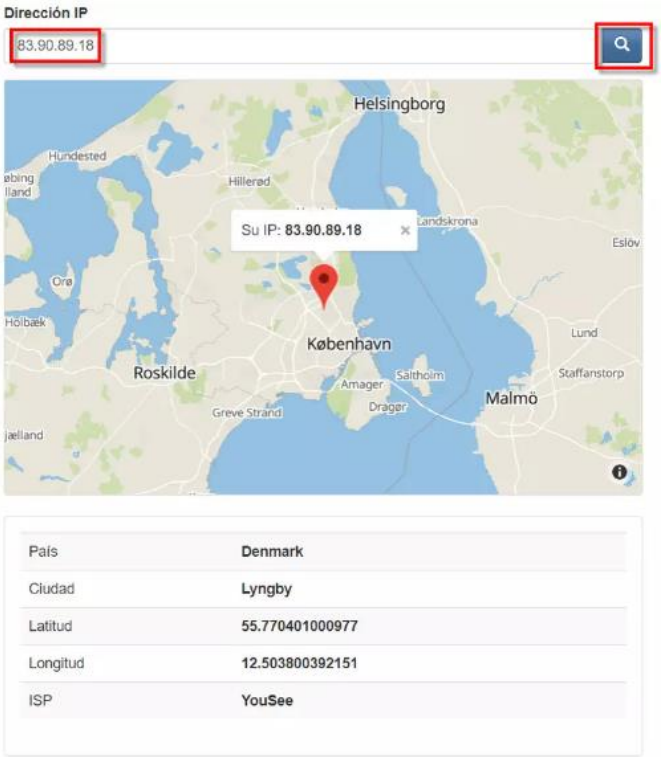


Figura 5-2 Ejemplo de información proporcionada por una dirección IP [178]

- **Proveedor de servicios de Internet (ISP):** la IP puede revelar el proveedor de servicios de Internet que suministra la conexión a la red del dispositivo. Esto da una idea de la empresa que proporciona el acceso a Internet.
- **Tipo de conexión:** la dirección IP puede indicar si el dispositivo está conectado a través de una red doméstica, móvil o empresarial.

- **Información sobre el dominio:** a veces, se puede realizar una búsqueda inversa de IP para obtener información sobre el dominio asociado con esa dirección IP.
- **Tipo de dispositivo:** información sobre el software que está utilizando, como el sistema operativo y el navegador.

Figura 5-3 Ejemplo de información proporcionada por una dirección IP

- **Mediante un escaneo de puertos se puede analizar qué servicios están activos.**

En cualquier caso y siguiendo diferentes técnicas, se puede llegar a monitorizar el uso que se hace por parte de quien utiliza una dirección para acceder a Internet y conocer cosas como el nombre del usuario, su teléfono o correo electrónico. De hecho, la Canadian Privacy Commissioner's Office, realizó una prueba de concepto, obteniendo la siguiente información de una IP:

- Afiliación religiosa.
- Estado físico.
- Fotos compartidas.
- Revisiones de la Wikipedia.
- Información relacionada con un problema legal de un usuario.

Siguiendo con el ejemplo práctico, se identificó y aisló la dirección IP de la persona que había editado la página web de Wikipedia, ya que las direcciones IP de este sitio son públicas. Al introducirla en un motor de búsqueda, se obtuvo información adicional, como otras ediciones realizadas en la enciclopedia en línea y hasta las visitas que había realizado en un determinado foro, revelando aspectos de su orientación sexual [179].

5.1.4 Información que un usuario proporciona voluntariamente al interactuar con diferentes sitios web.

Dentro de esta categoría la mayor cantidad de información es la que se genera en las redes sociales. De esta manera podemos acceder a los perfiles públicos de los usuarios en redes sociales, así como ver toda la información accesible.

Por otro lado, los "me gusta" o *likes*, las retransmisiones de contenido, los amigos, los seguidores las publicaciones, las fotos, etc, proporcionan una gran cantidad de información de un usuario.

Otra fuente de información son las reseñas que hacen los usuarios en Google o en las diferentes páginas web donde valoran la atención o la calidad de los productos que se han comprado o al rellenar

cualquier formulario del sitio, como por ejemplo en Newsletter, Blog, Eventos, Patrocinios, Proyectos, etc.

Con los registros personales dentro de las páginas web para poder realizar compras, es otro de los casos en los que se proporciona información de forma voluntaria.

5.1.5 Otros datos recopilados automáticamente mientras el usuario está navegando por Internet

Otros datos recopilados mientras el usuario navega por Internet son la forma en que los visitantes interactúan con un sitio, las fuentes de tráfico, visitantes únicos, recorrido por el sitio; se puede obtener una visión general de tu tráfico a diario, semanal, mensual o anual; rastrear qué sitios están enviando tráfico a nuestro sitio web; averiguar qué páginas de tu sitio web son más populares; registrar con qué palabras se hacen búsquedas en el sitio, etc.

5.2 Detección de la huella digital

Cuando se explora Internet, es posible dejar un rastro que podría poner en peligro nuestra privacidad. Esta información podría ser utilizada por terceros y, en algunos casos, incluso podrían aprovecharla para lanzar campañas de malware en nuestra contra. Es crucial tener en mente que la privacidad es uno de los aspectos más críticos que se deben proteger, aunque lamentablemente, no siempre es fácil hacerlo.

Para identificar el alcance de la huella digital que existe en Internet, existen diferentes herramientas y métodos que se pueden utilizar

- Buscar el nombre de una persona en los motores de búsqueda: esto dará una idea general de lo que hay en Internet sobre esa persona.

  Cuando se busque el nombre en los motores de búsqueda, utilizar una variedad de palabras clave y frases. Por ejemplo, se puede buscar el nombre completo, nombre de pila, apellido, nombre de usuario en redes sociales o dirección de correo electrónico.

  También se pueden utilizar operadores de búsqueda para afinar los resultados. Por ejemplo, puedes usar el operador "site:www.example.com" para buscar el nombre solo en el sitio web de example.com.

OPERADORES	EJEMPLOS	DETALLES ADICIONALES
...""...	social media""	Te muestra los resultados de páginas y recursos que tienen con exactitud los términos que has puesto antes de las comillas.
OR	Instagram OR Youtube OR Snapchat	Te da cualquiera de los tres términos o palabras que has incluido en la búsqueda.
-	marketing digital -SEO	Hace una exclusión de cualquier palabra clave o término que escribas después de el guión.
*	* La escuela de negocios que arrasa entre los que no buscaban un máster	El asterisco hará que todas las palabras que has introducido en el buscador aparezcan en los resultados de tu búsqueda.
#..#	Instagram actualizaciones #2015..2022#	Busca la información en el periodo de tiempo que incluyes entre los numerales.
SITE:	site:thepowermba.com/es negocios digitales	Hace la búsqueda de la palabra clave solo en el sitio web que has ingresado después del comando.
()	(másters" OR "masters en línea") - cursos de Google"	Te da la posibilidad de hacer la combinación de diferentes comandos de búsqueda. En este caso verás que tu búsqueda se centra en másters o másters en línea a excepción de los cursos gratuitos de Google.

OPERADORES	EJEMPLOS	DETALLES ADICIONALES
INTITLE:	intitle:community management"	Te da resultados de páginas con esa palabra clave o frase exacta en el título.
ALLINTITLE:	intitle:comentarios redes sociales	Muestra resultados con cualquiera de esas palabras clave en el título.
INURL:	inurl:posicionamiento SEO	Ofrece resultados con esa palabra clave exacta en la URL.
ALLINURL:	allinurl:facebook youtube estrategia	Te muestra URLS de cualquiera de las palabras clave que ingreses.
INTEXT:	intext:cursos en línea	Te muestra los resultados de páginas o contenido que contenga la palabra clave que has introducido en cualquier parte del cuerpo del texto.
ALLINTEXT:	allintext:editores de imágenes	Te muestra como resultado aquellas páginas que contengan cualquiera de las palabras clave que has introducido en cualquier parte del cuerpo del texto.
FILETYPE:	filetype:pdf contenido digital	Te muestra resultados en PDF que contengan exclusivamente la palabra clave que has introducido.
RELATED:	related:facebook.com	Te muestra los resultados relacionados a esa URL.
CACHE:	cache:thepowermba.com/es	Te mostrará cuál fue la última página que ha cacheado Google de esa web.

Figura 5-4 Ejemplo operadores de búsqueda [180]

- Revisar las redes sociales: las cuentas de las redes sociales probablemente tengan información personal de los individuos publicada en ellas.

- Consultar los registros públicos: en algunos casos, es posible encontrar información personal en registros públicos, como los registros de propiedad o los registros electorales.

  En algunos países, es posible encontrar información personal en registros públicos, como los registros de propiedad o los registros electorales. Para consultar estos registros, es posible que se necesite solicitarlos a las autoridades competentes.

- Utilizar herramientas de búsqueda de información privada: hay varias herramientas disponibles en línea que pueden ayudar a encontrar información privada sobre una persona.

  Estas herramientas utilizan una variedad de métodos para recopilar información, como la búsqueda en motores de búsqueda, la exploración de redes sociales, y la consulta de registros públicos.

  Algunos ejemplos de herramientas de búsqueda de información privada incluyen [181] [182]:

  - **Whois:** esta herramienta permite encontrar información sobre el propietario de un dominio web.
  - **Pinterest:** esta herramienta permite encontrar imágenes que contengan tu nombre o tu rostro.
  - **Google Alerts:** esta herramienta permite recibir notificaciones cuando se publica información nueva sobre ti en Internet.
  - **Pipl:** es un motor de búsqueda diseñado para encontrar información de personas. Puede proporcionar datos públicos de diferentes fuentes en línea.
  - **Spokeo:** es una herramienta que permite buscar información pública, incluyendo perfiles de redes sociales, números de teléfono y direcciones.
  - **ZabaSearch:** es un motor de búsqueda de información pública que incluye registros de directorios, direcciones y números de teléfono.
  - **Snitch.Name:** una herramienta especializada en la búsqueda de información en redes sociales y otras bases de datos.
  - **Hunter.io:** una opción destacada para localizar a personas con fines laborales, especializada en identificar individuos a través de sus direcciones de correo electrónico.
  - **Webmii:** un buscador que utiliza el nombre y apellido de una persona para localizarla en diversas fuentes, proporcionando información integral sobre la misma.

Egosurfing:

Egosurfing es la práctica de buscar información sobre uno mismo en Internet.

Hay varias razones por las que las personas pueden hacer egosurfing. Algunas personas lo hacen simplemente por curiosidad, para ver qué información hay sobre ellas en Internet. Otros lo hacen para proteger su privacidad, para identificar información que no debería estar publicada y que quieren que sea eliminada. También hay personas que utilizan el egosurfing para promocionarse a sí mismas o para mejorar su imagen pública.

El egosurfing puede ser una herramienta útil para gestionar tu huella digital. Sin embargo, es importante tener en cuenta que no toda la información que se encuentra en Internet sobre ti es precisa o completa. Por lo tanto, es importante ser crítico con la información que encuentras y tomar medidas para proteger tu privacidad si es necesario.

Utilizar buscadores para comprobar tu huella digital

Una de las maneras de comprobar tu huella digital en línea es utilizando buscadores y motores de búsqueda para obtener una visión general de la información que está disponible públicamente sobre una persona. Para ello podemos realizar las siguientes actividades.

- Mediante el uso de del proyecto Panopticlick, de la EFF (Electronic Frontier Foundation), se puede realizar un test del navegador, y en pocos segundos mostrará la huella digital del navegador web.

https://coveryourtracks.eff.org/

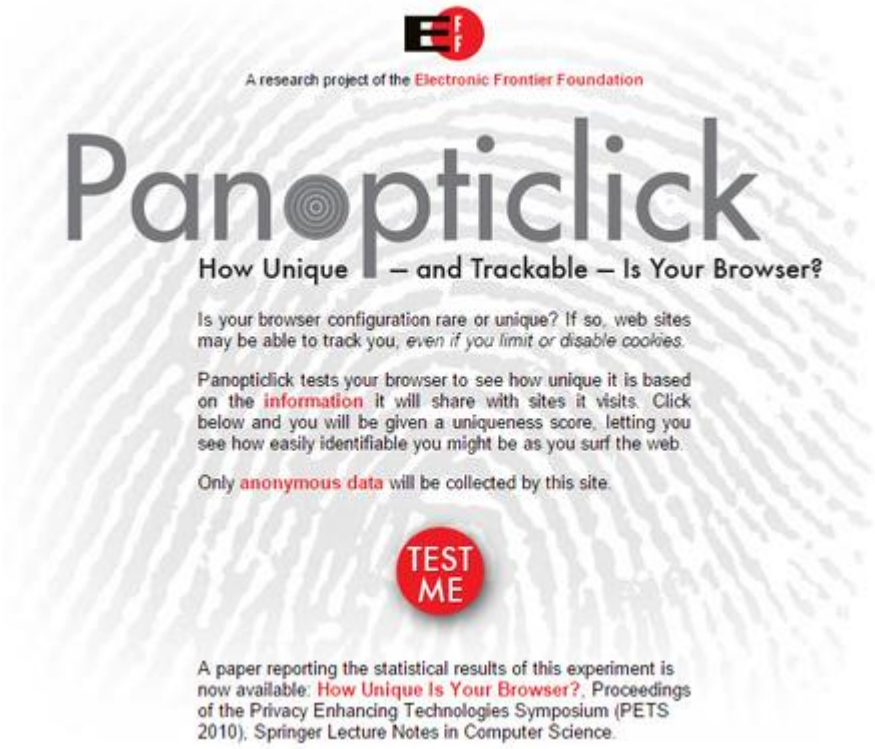


Figura 5-5 Proyecto Panopticlick [183]

Este recurso, sólo rastrea la actividad de un navegador web por cada test, con lo que se utilizan varios navegadores, habría que realizar el test en cada uno de los navegadores individualmente para identificar la información que dejan. Dicho proyecto también es válido para navegadores instalados en los teléfonos móviles.

- Introduce el nombre de la persona que se desea obtener información en los buscadores: es aconsejable utilizar diferentes buscadores (Google, Bing, Yahoo, Lycos, Ecosia, DuckDuckGo, Yandex, Baidu, Ask, etc [184]), ya que los resultados obtenidos serán diferentes.
    - Incluir el nombre y apellido y cualquier variación en la ortografía: si se cambia nombre, buscar tanto el nombre actual como el anterior. Revisar los resultados del buscador dará una idea de qué información sobre esa persona está disponible públicamente. Si alguno de los resultados se muestra de forma negativa, la persona se puede comunicar con el administrador del sitio para solicitar la eliminación.
    - Colocar el nombre entre comillas ("El Nombre") para buscar resultados específicos que contengan el nombre completo en ese orden.

- Usar la Búsqueda Avanzada utilizando funciones de búsqueda avanzada para refinar los resultados. Por ejemplo, se puede buscar un nombre junto con la ciudad en la que se vive, ocupación, o cualquier otro término relevante.

- Configurar Google Alerts o servicios similares (Talkwalker Alerts, Mention, Brand24, Awario, Social Mention, Trackur, Keyhole, etc [185] [186], [187]) que notifiquen cuando un nombre aparece en nuevos resultados en línea.

Buscar en otras de fuentes de información

Sitios como los portales inmobiliarios, plataformas como whitepages.com, registros médicos podrían poseer más datos sobre una persona de los deseados. Estos sitios suelen contener información personal, como número de teléfono, dirección y edad. Para eliminar esta información hay que ponerse en contacto con los sitios web y pedir la eliminación de la misma [188].

Auditar las cuentas

Durante la búsqueda de un nombre, es posible encontrar antiguas cuentas de redes sociales, publicaciones con bromas insensibles y obsoletas, o entradas de blog vergonzosas en las que se compartió demasiado de la vida personal.

5.3 Uso de la huella digital por terceros

En el ámbito de la publicidad y el seguimiento en línea, tres actores principales colaboran para rastrear a los usuarios y construir perfiles compuestos: los anunciantes, los agregadores y los editores [189].

- **Editore**s: se refiere a las empresas que difunden anuncios en línea mediante la combinación de anuncios con el contenido de diversas páginas web, juegos, entre otros.

- **Anunciantes:** hace referencia a las empresas que promocionan productos y servicios, siendo aquellas que buscan vender algo. En muchos casos, los anunciantes colaboran directamente con los editores.

  Cuando los anunciantes buscan una mayor especificidad, surgen los "agregadores" y brokers de datos, como las empresas BlueKai, Gravity, Rio, OutBrain y Dataium.

- **Los agregadores de datos** (presuntamente) recopilan información anónima de sus colaboradores y la utilizan para orientar los anuncios hacia el público más adecuado.

Por otro lado, el uso de la huella digital por terceros puede abordarse desde diferentes perspectivas, y a menudo está relacionado con la recopilación y el análisis de datos digitales:

- **Publicidad personalizada:** los anunciantes pueden utilizar la huella digital para mostrar anuncios adaptados a los intereses específicos de un usuario.

- **Seguimiento en línea:** los rastreadores pueden emplear la huella digital para monitorizar tu actividad en la web.

- **Ciberacoso:** los acosadores en línea pueden utilizar la huella digital para recopilar información sobre sus víctimas.

- **Robo de identidad:** los ladrones de identidad pueden aprovechar la huella digital para recopilar información que luego utilizan para llevar a cabo el robo de identidad.

- **Perfiles de Usuario:** plataformas de redes sociales y otros servicios en línea pueden crear perfiles detallados de usuarios basados en su huella digital, lo que puede incluir preferencias, conexiones sociales, historial de búsqueda, etc.

- **Investigación y Verificación:** empleadores, agencias de investigación o entidades gubernamentales pueden utilizar la huella digital para verificar la identidad de personas o realizar investigaciones, en algunos casos, con propósitos legítimos.

5.4 Métodos para eliminar la huella digital.

Eliminar completamente la huella digital en línea es un desafío, ya que la información puede estar dispersa en varios lugares y algunos datos pueden ser difíciles de eliminar por completo. Sin embargo, existen métodos y prácticas que pueden ayudar a reducir la visibilidad de la huella digital y minimizar la exposición en línea.

Básicamente, las acciones a realizar para eliminar tu huella digital son las siguientes

Identificar el alcance de la huella digital:

Se recomienda comenzar por realizar un inventario de todas las cuentas y servicios en línea que se han utilizado en los últimos años, así como en todos los sitios en los que una persona se ha registrado. Esto incluye redes sociales, cuentas de correo electrónico, cuentas de compras en línea, etc.

Por lo tanto, se debe rastrear Internet en busca de información personal o sensible en los diferentes medios de comunicación, boletines oficiales, boletines mercantiles, redes sociales, foros, artículos y blogs, buscadores de Internet, datos de directorios, bases de datos y páginas de información empresarial, antecedentes penales, listas de morosos, reclamaciones judiciales, publicaciones de imágenes o videos, datos médicos, opiniones, reseñas, etc.

Una forma un tanto ortodoxa de localizar cuentas "olvidadas" es comprobar las contraseñas que se han filtrado en Internet, actividad que nos puede orientar. Para ello se puede utilizar por ejemplo el siguiente recurso:

https://haveibeenpwned.com/

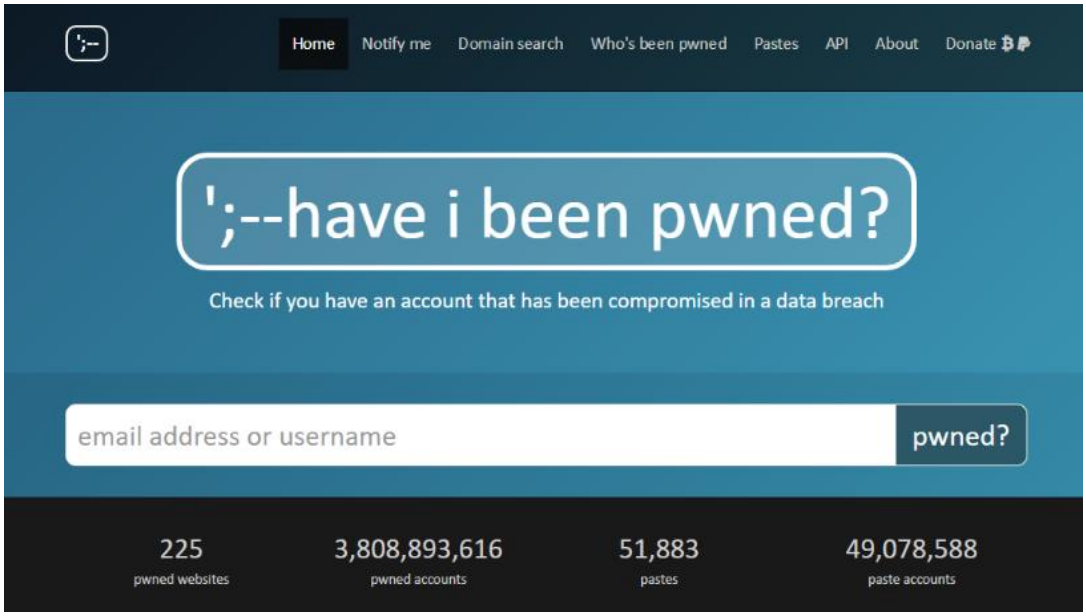


Figura 5-6 Pagina web para comprobar contraseñas comprometidas [190]

Cortar la difusión de datos en la fuente

Borrar la información o publicaciones innecesarias de las redes sociales, editar los perfiles de dichas redes aumentado los niveles de seguridad y privacidad (por ejemplo, modificando el nombre, nombre de usuario y descripción) o incluso eliminando las cuentas de las redes sociales, foros o webs. Toda información publicada en redes sociales puede ser indexada y aparecer en Internet usando los buscadores [191].

Dentro de las redes sociales, podemos encontrar formularios de "reporte de derechos de autor" para solicitar la eliminación de información privada compartida sobre un individuo por terceros. También se pueden reportar infracciones de privacidad de una foto o video al propietario o equipo de administración de una red social.

Analizar la información personal de los diferentes grandes servicios web (ej. Microsoft, Google, Apple, Amazon, etc), tiendas online y servicios digitales, ya que pueden crear perfiles personales y valorar que información debe aparecer y que información debe ser borrada.

También existe la posibilidad de contactar con los webmasters de las páginas web a través de la sección "Contacto" de dichas páginas independientemente de que se tenga o no cuenta en dicha web para solicitar la eliminación de contenido.

Solicitar a los navegadores, amparándose en el cumplimiento de la ley europea de "Derecho al Olvido (RGPD)" que eliminen los datos del buscador. Esto lo hacen desindexando la información que no se desea que se comparta de los buscadores, la información/web no se elimina, pero no resulta visible a los buscadores. En función del lugar del mundo en el que nos encontremos, las leyes de RGPD tienen aplicaciones diferentes, esto significa, por ejemplo, que una información eliminada en Europa, puede permanecer visible en otra parte del mundo. Los buscadores analizan las solicitudes de eliminación de información realizadas y deciden la conveniencia del borrado de dicha información. En el caso de Google, o servicios de Google, existe un formulario (https://reportcontent.google.com/forms/rtbf) para realizar dichas solicitudes.

Otro recurso se puede intentar llevar a cabo acciones legales contra los sitios que no han eliminado la información sensible, pero es un procedimiento largo y costoso.

Como último recurso, podemos contratar una empresa especializada en eliminación de huella digital.

Realizar una limpieza digital profunda

Después de identificar el alcance de la huella digital y cortar la difusión de datos en la fuente, la siguiente acción a realizar es acometer una limpieza digital profunda de los dispositivos electrónicos. Para ello se pueden realizar las siguientes acciones:

- Evitar el rastreo del historial de navegación: de esta manera se mejora la privacidad y aumenta el rendimiento del navegador. Esto se puede conseguir con la utilización del modo incognito o navegación privada de los navegadores.
- Borrar las cookies y el historial de navegación de los navegadores.
- Limpiar los ordenadores personales: desinstalando los programas que ya no se utilicen, además se deben borrar los archivos temporales frecuentemente, los archivos duplicados, imágenes de baja resolución.
- Limpiar los dispositivos móviles: borrar la memoria caché y los datos temporales, las cookies y los datos del historial, desinstalar las aplicaciones innecesarias, limpiar las descargas eliminando los archivos que no se necesiten. En algún momento se puede considerar revertir el dispositivo a su estado de fábrica.
- Elimina los metadatos de los archivos antes de compartir imágenes o documentos en la web. Los metadatos pueden contener información sensible, como ubicaciones.
- Revisar todos los contactos tanto en las cuentas de correo como en las redes sociales y eliminar aquellos contactos que no se deban o deseen tener.

Adoptar un enfoque más consciente.

En la parte de adopción de un enfoque más consciente, se hace referencia a la aplicación del sentido común cuando se estén realizando actividades en línea y antes de publicar ninguna foto, video o realizar comentarios en redes sociales o blogs.

Considerar utilizar herramientas y servicios que respeten la privacidad.

Otra de las acciones a realizar es configurar las alertas online. Para ello existen diferentes alternativas, siendo las alertas de Google las más conocidas. Otras alertas online [192][193] que podemos configurar son TalkWalker Alerts, Awario, Mention, Mediatoolkit, Infoxicate, Social Mention, Newsly, alertas de Bing, Yahoo Alerts [194], etc. Este sistema de alertas es capaz de detectar e informar si un contenido o una mención a una persona aparece online.

DeGoogling

El movimiento DeGoogle es una iniciativa surgida e impulsada por defensores de la privacidad. Esta iniciativa se basa en la idea de que Google recopila una cantidad excesiva de datos personales de sus usuarios, que luego utiliza para fines comerciales.

El término "DeGoogle" se refiere a reducir la dependencia de Google para proteger la privacidad de una persona o de una empresa. Implica utilizar alternativas centradas en la privacidad en lugar de los servicios de Google.

La participación en el mercado en constante aumento del gigante de Internet otorga a la empresa un poder monopolístico en los espacios digitales. Esto dificulta encontrar alternativas viables a los productos de Google. Sin embargo, un número cada vez mayor de proyectos están trabajando para desarrollar alternativas que sean más respetuosas con la privacidad.

El movimiento DeGoogle se puede interpretar como parte de una oposición más amplia hacia las grandes empresas tecnológicas. La creciente conciencia sobre las implicaciones de privacidad y el deseo de reducir la dependencia de un único proveedor tecnológico están impulsando a las personas a explorar opciones más diversificadas en el panorama digital.

El proceso para eliminar Google pasa por identificar los servicios que proporciona o colaboran con Google y buscar servicios alternativos que se centren en la privacidad. Este proceso pasa eliminar los servicios de Google de los ordenadores personales y de los móviles.

Algunos de los servicios que Google proporciona son los siguientes [195] [196]:

- **Correo electrónico**: GMail
- **Contactos**: integrados con los servicios de las grandes compañías tecnológicas.
- **Calendario**: Google Calendar
- **Herramientas colaborativas**: Office, Teams, Zoom, Onedrive, Google Drive
- **Mensajería**: Facebook, Messenger, WhatsApp, Snapchat
- **Navegadores**: Chrome
- **Redes Sociales**: Facebook, Twitter, LinkedIn
- **OS**: Windows, Android, etc.
- **DNS**: 8.8.8.8

En las siguientes direcciones web, se pueden identificar soluciones alternativas a los servicios de Google centradas en la privacidad y la seguridad.

https://www.privacytools.io/ https://www.privacyguides.org/es/

En general, en términos de privacidad, se recomienda el uso de soluciones "*Open Source*" y el uso de protocolos estándar evitando las soluciones y protocolos propietarios y monitorizando de forma continua las herramientas utilizadas para identificar si mantienen o cambian sus términos de uso referentes a la privacidad.

Derecho al Olvido

El "derecho al olvido" en Internet es un concepto relacionado con la privacidad y la protección de datos personales. Se refiere al derecho que tiene una persona a solicitar la eliminación de información personal sobre sí misma que se encuentra en Internet, especialmente cuando dicha información es obsoleta, inexacta, o ya no es relevante. Este derecho busca permitir que las personas controlen su propia información personal y eviten que información desactualizada o perjudicial afecte injustamente su reputación.

El derecho al olvido se ha vuelto más relevante con la proliferación de motores de búsqueda y redes sociales, ya que la información personal puede ser fácilmente accesible en línea durante períodos prolongados. Este concepto se reconoció inicialmente en la Unión Europea, donde el Tribunal de Justicia de la Unión Europea estableció en 2014 que los motores de búsqueda, como Google, debían permitir a los individuos solicitar la eliminación de enlaces a información personal irrelevante o inexacta [197].

Es importante señalar que el derecho al olvido debe equilibrarse con otros derechos, como la libertad de expresión y el acceso a la información. La implementación y aplicación de este derecho varían según la jurisdicción y las leyes específicas de privacidad de cada país.

5.5 Desafíos y riesgos asociados.

La eliminación de la huella digital, entendida como la reducción o eliminación de la presencia digital de una persona en Internet, presenta desafíos y riesgos que deben considerarse cuidadosamente:

- **Persistencia de Datos:** Aunque se intente eliminar la huella digital, algunos datos pueden persistir en la red. Copias de información personal podrían estar almacenadas en servidores, bases de datos o en cachés de motores de búsqueda.
- **Dificultad en la Eliminación Total:** Es difícil lograr una eliminación total. Incluso si se eliminan perfiles en redes sociales o sitios web, es posible que haya copias de dicha información en otros lugares, y los motores de búsqueda pueden conservar enlaces antiguos durante algún tiempo.
- **Impacto en la Reputación Online:** La eliminación de la huella digital podría afectar a la presencia en línea de una persona, especialmente si se trata de información positiva o relacionada con la reputación profesional.
- **Desconexión Social:** Al eliminar perfiles en redes sociales o reducir la presencia en línea, se puede experimentar una desconexión social, ya que muchas personas utilizan estas plataformas para mantenerse en contacto.
- **Riesgo de Pérdida de Oportunidades:** En algunos casos, una presencia en línea puede ser necesaria o beneficiosa para oportunidades laborales, educativas o sociales. Eliminar completamente la huella digital puede limitar estas oportunidades.
- **Nuevos Datos Generados Constantemente:** Aunque se eliminen los datos existentes, es probable que nuevos datos se generen continuamente a medida que un individuo interactúa en línea, lo que puede dificultar los esfuerzos de eliminar completamente la huella digital.
- **Problemas Legales y Éticos:** Algunas acciones para eliminar la huella digital podrían entrar en conflicto con las leyes de privacidad o términos de servicio de plataformas en línea, lo que podría tener consecuencias legales o éticas.

- **Riesgos de Seguridad:** Al borrar cuentas o información, existe el riesgo de que otras personas intenten suplantar una identidad o utilizar la información que se ha eliminado de manera malintencionada.

La eliminación de la huella digital es un proceso complejo que debe abordarse con precaución. Antes de tomar medidas drásticas, es aconsejable evaluar los posibles riesgos y beneficios, considerando cuidadosamente cómo afectará a la vida personal, profesional y social.

6 CONCLUSIONES

6.1 Resumen de los hallazgos clave.

Después de terminar y revisar en detalle este trabajo, se extraen conclusiones fundamentales que han sido reiteradas a lo largo de los diversos capítulos tratados. La privacidad y seguridad de las redes no son cuestiones triviales; sino que más bien, deben ser consideradas como prioridades fundamentales en cualquier contexto digital.

Independientemente de las técnicas y metodologías empleadas, es crucial reconocer que la consecución de un anonimato y seguridad absolutos resulta inalcanzable. La selección y adaptación cuidadosa de las soluciones implementadas deben estar en sintonía con las necesidades, requisitos y objetivos específicos que se buscan lograr. En este sentido, se subraya la importancia de ejercer el sentido común y la precaución en cualquier entorno en el que nos encontremos. Esto se debe a que el nivel de anonimato y privacidad estará intrínsecamente limitado por la forma en que utilicemos las diferentes soluciones, y es esencial tener presente que la fortaleza de la seguridad de un sistema o entorno es tan sólida como el eslabón más débil de su cadena.

Durante la investigación, se ha identificado una diversidad de soluciones adaptadas a distintos casos de uso, y se ha observado que el estado del arte continúa en constante evolución. Numerosos expertos en diversas disciplinas se dedican incansablemente a explorar y desarrollar nuevas soluciones para abordar los desafíos emergentes en materia de privacidad y seguridad.

En el desarrollo del trabajo, hemos mencionado y explicado con mayor o menor acierto multitud de conceptos, muchos de los cuales son de actualidad y se emplean en diferentes ámbitos diariamente y otros son al contario los embriones o primeros pasos o estudios realizados para conseguir algunos objetivos relacionados con la temática que se ha tratado y que a su vez han servido de base para el estudio y posterior desarrollo de otros conceptos o soluciones.

Dentro del catálogo de mejores prácticas, se han propuesto multitud de medidas para adoptar. Sin embargo, se reconoce que la implementación completa de todas estas medidas es una tarea que demanda un conocimiento elevado, tanto en sistemas como en redes de comunicación. A pesar de ello, se destaca que muchos de estos enfoques están al alcance de usuarios con conocimientos intermedios, permitiendo su aplicación y utilización para fortalecer la seguridad y privacidad en un entorno digital. La clave radica en la conciencia y comprensión de las herramientas disponibles, así como en su aplicación de manera informada y adaptada a las circunstancias particulares.

6.2 Contribuciones al campo de estudio.

El ámbito de estudio abordado en este trabajo es notablemente extenso, abarcando una diversidad de aspectos que, lamentablemente, no es viable explorar en profundidad dentro de los límites establecidos para este trabajo. En respuesta a esta complejidad, hemos optado por realizar una compilación exhaustiva, recopilando información de diversas fuentes especializadas. Este enfoque nos ha permitido condensar en un solo documento una variedad de soluciones y orientaciones destinados a mejorar la conciencia de los usuarios y sus prácticas al configurar y navegar por Internet.

La esencia y valor añadido de este trabajo radica en proporcionar una referencia integral que sirva como guía para los usuarios en la mejora de su seguridad y privacidad en línea. Al reunir recomendaciones provenientes de diversas fuentes, hemos creado un recurso que aborda la complejidad del entorno digital de manera holística. Este documento no solo ofrece soluciones específicas, sino que también busca cultivar una mentalidad informada y proactiva en los usuarios.

Al aplicar las recomendaciones detalladas en este documento, un usuario adquiere la capacidad de identificar amenazas potenciales asociadas con la privacidad y la seguridad en línea. Además, se le brinda la oportunidad de explorar diversas aproximaciones para utilizar la web y las redes de comunicación de manera segura y, al mismo tiempo, con ciertas garantías de anonimato y ocultación. Este enfoque integral no solo capacita al usuario para tomar decisiones informadas, sino que también fomenta una participación consciente en la protección de su propia seguridad digital. Con este trabajo, aspiramos a ofrecer una herramienta valiosa que inspire prácticas responsables y seguras en el entorno cada vez más complejo de Internet y las comunicaciones digitales.

6.3 Limitaciones del trabajo y áreas para futuras investigaciones.

Durante la ejecución de esta investigación, nos hemos enfrentado a diversas limitaciones que hemos abordado con el máximo esfuerzo. Una de las dificultades notables ha sido la restricción de tiempo disponible y la necesidad de definir un enfoque para llevar a cabo el trabajo. La temática de la anonimización, ocultación y eliminación de la huella digital abarca un espectro de materias extraordinariamente amplio, tanto que cada una podría constituir la base para múltiples trabajos individuales. Como acabamos de mencionar, la amplitud de la temática permite un desarrollo más extenso y profundo de cada capítulo del trabajo.

Otra limitación que hemos enfrentado ha sido la escasez de tiempo y recursos para llevar a cabo demostraciones prácticas que ilustren las instalaciones, configuraciones y evaluaciones ("*benchmarks*") de las diversas soluciones propuestas. Este impedimento se atribuye principalmente a la carencia de una infraestructura adecuada que respalde el desarrollo y la prueba de las diferentes soluciones tratadas en el trabajo.

Finalmente, es importante destacar que la mayoría de las soluciones o propuestas presentadas a lo largo de este trabajo se basan en la experiencia personal del autor y en la recopilación de información proveniente de las diversas fuentes referenciadas durante la elaboración del trabajo. A pesar de las limitaciones mencionadas, se ha hecho un esfuerzo continuo para ofrecer una perspectiva informada y respaldada por la mejor información disponible en el tiempo y marco de referencia establecidos para este proyecto.

Por otro lado, y en referencia a las áreas de trabajo para futuras investigaciones podemos afirmar que, con el continuo avance tecnológico, las empresas e instituciones y las personas se encuentran ante constantes desafíos en el ámbito de la ciberseguridad. La naturaleza dinámica del entorno digital, que experimenta cambios significativos a diario, implica que las estrategias de ataque evolucionen al mismo ritmo y nos obliguen a estar continuamente estudiando la evolución de las diferentes técnicas y tecnologías para poder mantener la seguridad y el anonimato en el uso de Internet. No obstante y

viendo las últimas evoluciones y direcciones de desarrollo de la tecnología podemos definir como áreas para futuras investigaciones principalmente las tecnologías emergentes y disruptivas (EDTs[12]).

La comunidad científica mira expectante el desarrollo y evolución de estas tecnologías que no solo están cambiando industrias específicas, sino que también están generando nuevas oportunidades y desafíos que requieren adaptabilidad y una comprensión profunda de su impacto en la sociedad y la economía. Dentro de estas tecnologías o áreas de trabajo futuro podemos mencionar

- **Inteligencia Artificial (IA) y Aprendizaje Automático (ML):** La capacidad de las máquinas para aprender y tomar decisiones de manera autónoma ha transformado industrias como la salud, finanzas y manufactura.
- **Federated Learning:** que es un tipo de metodología que permite usar datos anonimizados en AI.
- **Blockchain y Criptomonedas:** La tecnología blockchain, conocida por ser la base de las criptomonedas como Bitcoin y Ethereum, ha alterado los paradigmas de seguridad y transparencia en las transacciones financieras y la gestión de datos.
- **Internet de las Cosas (IoT):** La interconexión de dispositivos cotidianos a través de Internet ha revolucionado sectores como la domótica, la salud y la logística, permitiendo la recopilación de datos en tiempo real para una toma de decisiones más eficiente.
- **Computación Cuántica:** Aunque aún está en sus primeras etapas, la computación cuántica tiene el potencial de transformar radicalmente la capacidad de procesamiento, afectando áreas como la criptografía y la simulación de sistemas complejos.
- **Realidad Aumentada (AR) y Realidad Virtual (VR):** Estas tecnologías han cambiado la forma en que experimentamos la información y el entretenimiento, con aplicaciones en la educación, la medicina y el diseño.
- **Biología Sintética y Edición Genética:** La capacidad de manipular genomas y crear organismos sintéticos tiene implicaciones significativas en la medicina, la agricultura y la producción de alimentos.
- **Vehículos Autónomos:** La automatización en el transporte, como los vehículos autónomos, está cambiando la forma en que nos movemos y plantea desafíos y oportunidades en la movilidad urbana.
- **Robótica Avanzada:** La robótica colaborativa y autónoma está siendo aplicada en la fabricación, la salud y la exploración espacial, entre otros campos.

Evidentemente, la seguridad, privacidad, anonimización, ocultación y eliminación de la huella digital en estos entornos es una de las tareas pendientes que se irán desarrollando conforme evolucionen las mencionadas tecnologías.

[12] EDT – Emerging and disruptive tecnologies

7 BIBLIOGRAFÍA

[1] "BOE-A-1978-31229 Constitución Española." Accessed: Sep. 24, 2023. [Online]. Available: https://www.boe.es/buscar/act.php?id=BOE-A-1978-31229

[2] "Carta de los Derechos Fundamentales de la Unión Europea," 2010.

[3] S. de España, "Protección de datos de carácter personal".

[4] "BOE-A-2018-16673 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales." Accessed: Sep. 24, 2023. [Online]. Available: https://boe.es/buscar/act.php?id=BOE-A-2018-16673

[5] "Importancia de la Ley de protección de datos en empresas | Grupo Atico34." Accessed: Sep. 24, 2023. [Online]. Available: https://protecciondatos-lopd.com/empresas/importancia-ley/

[6] "Data protection: why it matters and how to protect it - Access Now." Accessed: Sep. 24, 2023. [Online]. Available: https://www.accessnow.org/data-protection-matters-protect/

[7] L. F. Cranor, "Internet privacy," *Commun ACM*, vol. 42, no. 2, pp. 28–38, Feb. 1999, doi: 10.1145/293411.293440.

[8] "ACM: Digital Library: Communications of the ACM." Accessed: Sep. 24, 2023. [Online]. Available: https://dl.acm.org/doi/fullHtml/10.1145/293411.293440

[9] "¿Qué es una huella digital?" Accessed: Nov. 02, 2023. [Online]. Available: https://latam.kaspersky.com/resource-center/definitions/what-is-a-digital-footprint

[10] "¿Qué es la huella digital en internet? | Argentina.gob.ar." Accessed: Nov. 04, 2023. [Online]. Available: https://www.argentina.gob.ar/justicia/convosenlaweb/situaciones/que-es-la-huella-digital-en-internet

[11] "Huella digital en Internet: ¿Hasta dónde saben de mí?" Accessed: Nov. 03, 2023. [Online]. Available: https://www.inesem.es/revistadigital/informatica-y-tics/huella-digital-internet/

[12] A. Calderón Gómez, "Aprovechamiento de la huella digital: la opinión en las redes sociales," 2010. [Online]. Available: https://e-archivo.uc3m.es/handle/10016/10420

[13] "Cómo reducir tu huella digital - Veigler Formación." Accessed: Nov. 03, 2023. [Online]. Available: https://veiglerformacion.com/huella-digital-importancia-ejemplos/

[14] "La privacidad digital: Navegando seguros en la era digital." Accessed: Nov. 04, 2023. [Online]. Available: https://grupoadaptalia.es/blog/privacidad-digital-que-es/

[15] "La protección de datos en la UE." Accessed: Nov. 04, 2023. [Online]. Available: https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_es

[16] "Normas internacionales relativas a la privacidad digital | OHCHR." Accessed: Nov. 04, 2023. [Online]. Available: https://www.ohchr.org/es/privacy-in-the-digital-age/international-standards-relating-digital-privacy

[17] "Introducción a la anonimización de datos: Técnicas y casos prácticos | datos.gob.es." Accessed: Sep. 26, 2023. [Online]. Available: https://datos.gob.es/es/documentacion/introduccion-la-anonimizacion-de-datos-tecnicas-y-casos-practicos

[18] "¿Qué es la anonimización de datos? Protege tus datos." Accessed: Sep. 26, 2023. [Online]. Available: https://www.klippa.com/es/blog/informativo/anonimizacion-datos/

[19] "Introducción a la anonimización de datos: Técnicas y casos prácticos | datos.gob.es." Accessed: Sep. 25, 2023. [Online]. Available: https://datos.gob.es/es/documentacion/introduccion-la-anonimizacion-de-datos-tecnicas-y-casos-practicos

[20] D. Wang and A. Bakhai, "Clinical Trials - A Practical Guide to Design, Analysis, and Reporting," 2006.

[21] "4.1. Aleatorización | Unidad Reguladora y de Control de Datos Personales." Accessed: Sep. 26, 2023. [Online]. Available: https://www.gub.uy/unidad-reguladora-control-datos-personales/comunicacion/publicaciones/guia-criterios-disociacion-datos-personales/capitulo-4-conjunto-tecnicas

[22] "Aleatorización de bloques permutados en 2023 → STATOLOGOS®." Accessed: Sep. 26, 2023. [Online]. Available: https://statologos.com/aleatorizacion-de-bloques-permutados/?utm_content=cmp-true

[23] J. Bollo, S. Fernández-Ananin, and E. Targarona, "Ensayo clínico aleatorizado," *Cir Esp*, vol. 100, no. 7, pp. 442–444, jul. 2022, doi: 10.1016/J.CIRESP.2021.11.009.

[24] D. R. Taves, "Minimization: A new method of assigning patients to treatment and control groups," *Clin Pharmacol Ther*, vol. 15, no. 5, pp. 443–453, May 1974, doi: 10.1002/CPT1974155443.

[25] S. Estrada, M. Arancibia, J. Stojanova, and C. Papuzinski, "General concepts in biostatistics and clinical epidemiology: Experimental studies with randomized clinical trial design," *Medwave*, vol. 20, no. 3, 2020, doi: 10.5867/MEDWAVE.2020.02.7869.

[26] "Qué es la privacidad diferencial." Accessed: Sep. 25, 2023. [Online]. Available: https://www.manueldelgado.com/p/que-es-la-privacidad-diferencial

[27] "'Differential privacy': el sistema clave para garantizar la privacidad de datos." Accessed: Sep. 25, 2023. [Online]. Available: https://www.redseguridad.com/especialidades-tic/proteccion-de-datos/differential-privacy-el-sistema-clave-para-garantizar-la-privacidad-de-los-datos_20210222.html

[28] M. Arias and O. Sangrador, "Ensayo clínico (III). Aleatorización. Enmascaramiento".

[29] "¿Qué es la privacidad diferencial y cómo protege tus datos?" Accessed: Sep. 25, 2023. [Online]. Available: https://ciberseguridad.com/guias/prevencion-proteccion/privacidad-diferencial/

[30] "Privacidad diferencial - Glosario FineProxy." Accessed: Sep. 25, 2023. [Online]. Available: https://fineproxy.org/es/wiki/differential-privacy/

[31] "Privacidad Diferencial - El Guardián de tu Seguridad en la Red." Accessed: Sep. 26, 2023. [Online]. Available: https://aprendeinformaticas.com/privacidad-diferencial/

[32] B. Bebensee, "Local Differential Privacy: a tutorial", Accessed: Sep. 26, 2023. [Online]. Available: https://github.com/google/rappor

[33] "¿Qué es la privacidad diferencial y cómo protege tus datos?" Accessed: Sep. 26, 2023. [Online]. Available: https://ciberseguridad.com/guias/prevencion-proteccion/privacidad-diferencial/

[34] A. Cavoukian, "Information and Privacy Commissioner/Ontario Creation of a Global Privacy Standard".

[35] "Privacidad por diseño: principios y cómo implementarla." Accessed: Sep. 26, 2023. [Online]. Available: https://ciberseguridad.com/normativa/espana/medidas/privacidad-por-diseno/

[36] "Cómo la privacidad diferencial complementa a la anonimización para garantizar la seguridad de los datos." Accessed: Sep. 26, 2023. [Online]. Available: https://blog.pangeanic.com/es/privacidad-diferencial-y-anonimizacion

[37] L. Sweeney, "L. Sweeney. k-anonymity: a model for k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY 1," *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[38] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "ℓ-Diversity: Privacy Beyond k-Anonymity".

[39] K. A. Hua, F. Liu, and Y. Cai, "Query l-diversity in Location-Based Services," in *2013 IEEE 14th International Conference on Mobile Data Management*, Los Alamitos, CA, USA: IEEE Computer Society, May 2009, pp. 436–442. doi: 10.1109/MDM.2009.72.

[40] P. Vassiliadis, "Review: l-diversity-privacy beyond k-anonymity".

[41] N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and-Diversity".

[42] C. J. Hoofnagle, B. van der Sloot, and F. Z. Borgesius, "The European Union general data protection regulation: What it is and what it means," *Information and Communications Technology Law*, vol. 28, no. 1, pp. 65–98, Jan. 2019, doi: 10.1080/13600834.2019.1573501.

[43] "PII, PHI, PCI: What is the Difference? Audit Compliance." Accessed: Dec. 06, 2023. [Online]. Available: https://linfordco.com/blog/pii-phi-pci-differences/

[44] "Diferencia entre anonimización y seudonimización | Bufete Mas y Calvet. Abogados." Accessed: Nov. 05, 2023. [Online]. Available: https://mascalvet.com/diferencia-entre-anonimizacion-y-seudonimizacion/

[45] "Pseudonymization for health applications." Accessed: Nov. 05, 2023. [Online]. Available: https://www.chino.io/compliance/pseudonymization-for-health-applications

[46] van H.C.A. Tilborg and S. Jajodia, "Encyclopedia of Cryptography and Security," *Encyclopedia of Cryptography and Security*, 2011, doi: 10.1007/978-1-4419-5906-5.

[47] A. J. Menezes, P. C. van Oorschot, and S. a Vanstone, "Hash Functions and Data Integrity BT - Handbook of Applied Cryptography," *Handbook of Applied Cryptography*, no. 9, pp. 321–383, 1996, Accessed: Nov. 06, 2023. [Online]. Available: http://cacr.uwaterloo.ca/hac/about/chap9.pdf

[48] "Pseudonymisation techniques and best practices", doi: 10.2824/247711.

[49] M. Bellare, R. Canetti, and H. Krawczyk, "Keying Hash Functions for Message Authentication," in *Advances in Cryptology — CRYPTO '96*, N. Koblitz, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 1–15.

[50] "Cifrado y Privacidad III: Cifrado Homomórfico | AEPD." Accessed: Dec. 06, 2023. [Online]. Available: https://www.aepd.es/prensa-y-comunicacion/blog/cifrado-privacidad-iii-cifrado-homomorfico

[51] Aepd, "Guía básica de anonimización Elaborada por Autoridad Nacional de Protección de Datos de Singapur (PDPC-Personal Data Protection Commission Singapore)".

[52] R. Merkle, *A Digital Signature Based on a Conventional Encryption Function*, vol. 293. 1987. doi: 10.1007/3-540-48184-2_32.

[53] L. Lamport, "Password Authentication with Insecure Communication," *Commun. ACM*, vol. 24, no. 11, pp. 770–772, Nov. 1981, doi: 10.1145/358790.358797.

[54] B. H. Bloom, "Space/Time Trade-Offs in Hash Coding with Allowable Errors," *Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul. 1970, doi: 10.1145/362686.362692.

[55] S. G. Weber, "On Transaction Pseudonyms with Implicit Attributes," *Cryptology ePrint Archive*, 2012.

[56] H. Wu and F. Wang, "A survey of noninteractive zero knowledge proof system and its applications," *Scientific World Journal*, vol. 2014, 2014, doi: 10.1155/2014/560484.

[57] G. Danezis *et al.*, "Privacy and Data Protection by Design - from policy to engineering," Jan. 2015, doi: 10.2824/38623.

[58] "La importancia de la anonimización y la privacidad de datos | datos.gob.es." Accessed: Dec. 06, 2023. [Online]. Available: https://datos.gob.es/es/blog/la-importancia-de-la-anonimizacion-y-la-privacidad-de-datos

[59] N. P. Hoang and D. Pishva, *Anonymous Communication and its Importance in Social Networking*. 2014. doi: 10.1109/ICACT.2014.6778917.

[60] Y. Xia, R. Chen, J. Su, and H. Zou, "Balancing anonymity and resilience in anonymous communication networks," *Comput Secur*, vol. 101, p. 102106, 2021, doi: https://doi.org/10.1016/j.cose.2020.102106.

[61] eerdin, "Anonymous Communication Systems: Usage Analysis and Attack Mechanisms," 2012.

[62] P. Mittal and N. Borisov, "Information Leaks in Structured Peer-to-Peer Anonymous Communication Systems," *ACM Transactions on Information and System Security (TISSEC)*, vol. 15, no. 1, Mar. 2012, doi: 10.1145/2133375.2133380.

[63] S. J. Murdoch and R. N. M. Watson, "Metrics for Security and Performance in Low-Latency Anonymity Systems", Accessed: Oct. 04, 2023. [Online]. Available: http://www.cl.cam.ac.uk/users/{sjm217,rnw24}

[64] M. K. Reiter and A. D. Rubin, "Crowds: Anonymity for Web Transactions," *ACM Trans. Inf. Syst. Secur.*, vol. 1, no. 1, pp. 66–92, Nov. 1998, doi: 10.1145/290163.290168.

[65] S. J. Murdoch, "Quantifying and measuring anonymity," in *International Workshop on Data Privacy Management*, 2013, pp. 3–13.

[66] T. Lu, Z. Du, and Z. Wang, "A Survey on Measuring Anonymity in Anonymous Communication Systems," *IEEE Access*, vol. PP, p. 1, May 2019, doi: 10.1109/ACCESS.2019.2919322.

[67] D. Chaum and D. L. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," 1981.

[68] "The Dining Cryptographers Problem." Accessed: Oct. 03, 2023. [Online]. Available: https://www.cs.cornell.edu/people/egs/herbivore/dcnets.html

[69] J. Feigenbaum and B. Ford, "Seeking Anonymity in an Internet Panopticon," *Commun ACM*, vol. 58, Dec. 2013, doi: 10.1145/2714561.

[70] C. Iacomini, "Infraestructura de un sistema confidencial y anónimo de transporte con rutas dinámicas y aleatorias," jun. 2020. [Online]. Available: https://oa.upm.es/64125/

[71] "Nace PrivaTegrity, el sustituto de la red TOR y el VPN | Computer Hoy." Accessed: Oct. 03, 2023. [Online]. Available: https://computerhoy.com/noticias/software/nace-privategrity-sustituto-red-tor-vpn-38963

[72] "PrivaTegrity, une alternative à Tor vraiment anonyme." Accessed: Oct. 03, 2023. [Online]. Available: https://www.silicon.fr/un-gourou-du-chiffrement-lance-privategrity-une-alternative-a-tor-135394.html

[73] "The Father of Online Anonymity Has a Plan to End the Crypto War | WIRED." Accessed: Oct. 04, 2023. [Online]. Available: https://www.wired.com/2016/01/david-chaum-father-of-online-anonymity-plan-to-end-the-crypto-wars/

[74] "Onion Routing." Accessed: Oct. 03, 2023. [Online]. Available: https://www.onion-router.net/

[75] "Tor Project | Anonimato en línea." Accessed: Oct. 02, 2023. [Online]. Available: https://www.torproject.org/es/

[76] "RED TOR. ¿Qué es Cómo Funciona Cómo me conecto? ▷ Guía 2023." Accessed: Oct. 03, 2023. [Online]. Available: https://internetpasoapaso.com/red-tor/

[77] P. Syverson, G. Tsudik, M. Reed, and C. Landwehr, "Towards an Analysis of Onion Routing Security".

[78] R. Snader and N. Borisov, "A Tune-up for Tor: Improving Security and Performance in the Tor Network".

[79] "Red TOR: qué es, cómo funciona y cómo se usa." Accessed: Oct. 03, 2023. [Online]. Available: https://www.xataka.com/basics/red-tor-que-como-funciona-como-se-usa

[80] M. Parameswaran, A. Susarla, and A. Whinston, "P2P networking: An information-sharing alternative," *Computer (Long Beach Calif)*, vol. 34, pp. 31–38, Aug. 2001, doi: 10.1109/2.933501.

[81] *Informática peer-to-peer: tecnologías para compartir y colaborar en la red*. David; Hillsboro, Oregón: Intel Press, 2001.

[82] S. Saroiu, K. P. Gummadi, and S. D. Gribble, "Medición y análisis de las características de los hosts Napster y Gnutella," *Sistemas multimedia*, vol. 9, pp. 170–184.

[83] R. A B C D Steinmetz and K. Wehrle, *Apuntes de conferencias en Ciencias de la Computación*. Berlín, Heidelberg: Springer, 2005.

[84] A. Oram, *Peer-To-Peer: Harnessing the Benefits of a Disruptive Technologies*. Sebastopol CA: O'Reilly, 2001.

[85] A. Kobusińska, J. Brzeziński, M. Boroń, L. Inatlewski, M. Jabczyński, and M. MacIejewski, "A branch hash function as a method of message synchronization in anonymous P2P conversations," *International Journal of Applied Mathematics and Computer Science*, vol. 26, no. 2, pp. 479–493, Jun. 2016, doi: 10.1515/AMCS-2016-0034.

[86] R. Endsuleit and T. Mie, *Censorship-resistant and anonymous P2P filesharing*. 2006. doi: 10.1109/ARES.2006.41.

[87] R. P. M. D. D-, "Redes Peer to Peer y Tecnologia JXTA."

[88] "Métodos para compartir archivos y contenidos en Internet (III): P2P, el auténtico espíritu del file sharing." Accessed: Oct. 07, 2023. [Online]. Available: https://www.xatakamovil.com/conectividad/metodos-para-compartir-archivos-y-contenidos-en-internet-iiip2p-el-autentico-espiritu-del-file-sharing

[89] "Peer-to-Peer Network (P2P) - NETWORK ENCYCLOPEDIA." Accessed: Oct. 07, 2023. [Online]. Available: https://networkencyclopedia.com/peer-to-peer-network-p2p/?utm_content=cmp-true

[90] "De igual a igual Contenido y Desarrollo histórico." Accessed: Oct. 07, 2023. [Online]. Available: https://hmong.es/wiki/Peer-to-peer

[91] T. Koskela, O. Kassinen, E. Harjula, and M. Ylianttila, "P2P Group Management Systems: A Conceptual Analysis," *ACM Comput. Surv.*, vol. 45, no. 2, Mar. 2013, doi: 10.1145/2431211.2431219.

[92] W. Allasia, "INTEROPERABLE DIGITAL RIGHTS MANAGEMENT ON STRUCTURED PEER-TO-PEER OVERLAY NETWORKS," 2008. doi: 10.13140/RG.2.1.2376.1521.

[93] A. Arunachalam and V. Ravi, *A study of the resource discovery approaches in Mobile Peer-to-Peer Networks*. 2021. doi: 10.36227/techrxiv.13726114.

[94] "Arquitectura Peer to Peer (P2P)." Accessed: Oct. 07, 2023. [Online]. Available: https://reactiveprogramming.io/blog/es/estilos-arquitectonicos/p2p

[95] X. Shen, Heather, J. Buford, and M. Akon, *Handbook of Peer-to-Peer Networking*. Nueva York: Springer, 2009.

[96] "Tixati.com - Home." Accessed: Jan. 15, 2024. [Online]. Available: https://tixati.com/

[97] "Utilizar Kad con eMule | www.emule.es." Accessed: Jan. 15, 2024. [Online]. Available: https://www.emule.es/blog/emule/utilizar-kad-con-emule/

[98] "Home - YaCy." Accessed: Jan. 15, 2024. [Online]. Available: https://yacy.net/

[99] "Cómo utilizar la red de distribución de contenido de Coral (CoralCDN)." Accessed: Jan. 15, 2024. [Online]. Available: https://lanera-austral.com/es/how-to/388-how-to-use-coral-content-distribution-network-coralcdn.html

[100] D. Korzun and A. Gurtov, *Sistemas P2P estructurados: fundamentos de organización jerárquica, enrutamiento, escalado y seguridad*.

[101] R. Ranjan, Lipo, A. Harwood, Shanika, and R. Buyya, *Servicio de descubrimiento de recursos descentralizado para redes federadas a gran escala" (PDF). Archivado desde el original (PDF) el 10 de septiembre de*. 2008.

[102] "Conexiones Peer to Peer. P2P - Tech Riders." Accessed: Oct. 09, 2023. [Online]. Available: https://techriders.tajamar.es/conexiones-peer-to-peer-p2p/

[103] E. H. T. B. Brands and G. Karagiannis, "Taxonomy of P2P Applications," *2009 IEEE Globecom Workshops, Gc Workshops 2009*, pp. 1–8, Dec. 2009, doi: 10.1109/GLOCOMW.2009.5360707.

[104] "How the Old Napster Worked | HowStuffWorks." Accessed: Oct. 16, 2023. [Online]. Available: https://computer.howstuffworks.com/napster.htm

[105] P. Evangelista *et al.*, *EbitSim: An Enhanced BitTorrent Simulation Using OMNeT++ 4*. 2011. doi: 10.1109/MASCOTS.2011.46.

[106] M. Ripeanu, "Peer-to-peer architecture case study: Gnutella network," *Proceedings - 1st International Conference on Peer-to-Peer Computing, P2P 2001*, pp. 99–100, 2001, doi: 10.1109/P2P.2001.990433.

[107] "Freenet." Accessed: Oct. 16, 2023. [Online]. Available: https://staging.freenetproject.org/es/index.html

[108] "maxresdefault.jpg (1280×720)." Accessed: Oct. 16, 2023. [Online]. Available: https://i.ytimg.com/vi/yNi64lNobIU/maxresdefault.jpg

[109] B. Beverly Yang and H. Garcia-Molina, "Designing a super-peer network," in *Proceedings 19th International Conference on Data Engineering (Cat. No.03CH37405)*, 2003, pp. 49–60. doi: 10.1109/ICDE.2003.1260781.

[110] "skype-arquitectura-red.png (504×585)." Accessed: Oct. 16, 2023. [Online]. Available: https://baluart.net/files/images/2988/skype-arquitectura-red.png

[111] "What is I2P? Definition, uses, and pros and cons | NordVPN." Accessed: Dec. 02, 2023. [Online]. Available: https://nordvpn.com/es/blog/what-is-i2p/

[112] "I2P con 2015 - Growing the Network, Spreading the Word - YouTube." Accessed: Dec. 04, 2023. [Online]. Available: https://www.youtube.com/watch?v=_R85duCOWsY&t=221s

[113] "Intro - I2P." Accessed: Dec. 02, 2023. [Online]. Available: https://geti2p.net/en/about/intro

[114] "Introduction to Anonymizing Networks – Tor vs I2P | Infosec." Accessed: Dec. 02, 2023. [Online]. Available: https://resources.infosecinstitute.com/topics/general-security/anonymizing-networks-tor-vs-i2p/

[115] J. Kosiński, *DEEPWEB AND DARKNET – POLICE VIEW*. 2015.

[116] "Así es Freenet, deep web alternativa a Tor e I2P." Accessed: Dec. 04, 2023. [Online]. Available: https://www.genbeta.com/a-fondo/asi-es-freenet-deep-web-alternativa-a-tor-e-i2p

[117] "Hyphanet." Accessed: Dec. 04, 2023. [Online]. Available: https://www.hyphanet.org/index.html

[118] "Zeronet Wants to Replace the Dark Web by Marrying Bitcoin to Bittorrent Over Tor – Featured Bitcoin News." Accessed: Dec. 05, 2023. [Online]. Available: https://news.bitcoin.com/zeronet-replacing-dark-web-marrying-bitcoin-bittorrent-tor/

[119] Y. Han, D. Xu, J. Gao, and L. Zhu, "Using Blockchains for Censorship-Resistant Bootstrapping in Anonymity Networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13407 LNCS, pp. 240–260, 2022, doi: 10.1007/978-3-031-15777-6_14.

[120] "¿Qué es una huella digital?" Accessed: Sep. 30, 2023. [Online]. Available: https://latam.kaspersky.com/resource-center/definitions/what-is-a-digital-footprint

[121] "Nine Steps to Reduce Your Digital Footprint with Ease." Accessed: Sep. 30, 2023. [Online]. Available: https://hotbotvpn.com/blog/9-steps-to-reduce-your-digital-footprint/

[122] "Sistema operativo Live-USB: qué es y cómo ayuda a reparar nuestro PC." Accessed: Oct. 07, 2023. [Online]. Available: https://www.softzone.es/programas/sistema/sistema-operativo-live-usb-que-es/

[123] "Navegación Segura Y Privada Ti Y Tu Empresa Parte I | Empresas | INCIBE." Accessed: Sep. 30, 2023. [Online]. Available: https://www.incibe.es/empresas/blog/navegacion-segura-y-privada-ti-y-tu-empresa-parte-i

[124] "Consejos para reducir tu huella digital en Internet y protegerte de ciberdelitos | Mujer Ejecutiva." Accessed: Sep. 30, 2023. [Online]. Available: https://mundoejecutivo.com.mx/mujer-ejecutiva/consejos-para-reducir-tu-huella-digital-en-internet-y-protegerte-de-ciberdelitos/

[125] "HTTPS vs HTTP – Why You Need to Act." Accessed: Oct. 01, 2023. [Online]. Available: https://www.owltree.co.uk/articles-and-news/https-vs-http/

[126] "10 consejos útiles para navegar por internet de forma segura – La Neurona." Accessed: Oct. 01, 2023. [Online]. Available: https://laneurona.com/texto/10-consejos-utiles-para-navegar-seguro-por-internet/

[127] "Motivos para reducir tu huella digital (y 3 formas de hacerlo) - Avira Blog." Accessed: Sep. 30, 2023. [Online]. Available: https://www.avira.com/es/blog/motivos-para-reducir-tu-huella-digital-y-3-formas-de-hacerlo

[128] "Cómo proteger y reducir tu huella digital - State Farm®." Accessed: Sep. 30, 2023. [Online]. Available: https://es.statefarm.com/simple-insights/familia/como-reducir-y-proteger-tu-huella-digital

[129] "Browser Password Manager: Flawed Security | Fractional CISO." Accessed: Oct. 01, 2023. [Online]. Available: https://fractionalciso.com/browser-password-managers-flawed-security-by-design/

[130] "¿Son seguras las contraseñas almacenadas en navegadores? | Blog oficial de Kaspersky." Accessed: Oct. 01, 2023. [Online]. Available: https://www.kaspersky.es/blog/how-to-store-passwords-securely/29106/

[131] "How to Extract Chrome Passwords in Python? - GeeksforGeeks." Accessed: Oct. 01, 2023. [Online]. Available: https://www.geeksforgeeks.org/how-to-extract-chrome-passwords-in-python/

[132] "Extract stored Chrome Passwords and Decrypt them using Python - Geeky Humans." Accessed: Oct. 01, 2023. [Online]. Available: https://geekyhumans.com/extract-stored-chrome-passwords-and-decrypt-them-using-python/

[133] "How to Extract Chrome Passwords in Python - Python Code." Accessed: Oct. 01, 2023. [Online]. Available: https://thepythoncode.com/article/extract-chrome-passwords-python?utm_content=cmp-true

[134] "Registrarte Con Tu Cuenta De Google Facebook O Twitter Ventajas E | Ciudadanía | INCIBE." Accessed: Oct. 15, 2023. [Online]. Available: https://www.incibe.es/ciudadania/blog/registrarte-con-tu-cuenta-de-google-facebook-o-twitter-ventajas-e

[135] "Facebook." Accessed: Oct. 04, 2023. [Online]. Available: https://www.facebook.com/legal/terms_preview_DSA

[136] "Cómo enviar un e-mail anónimo e imposible de rastrear." Accessed: Oct. 07, 2023. [Online]. Available: https://www.redeszone.net/tutoriales/seguridad/como-enviar-email-anonimo/

[137] "Protección de la Navegación: Área de Sistemas de la Información y las Comunicaciones: UPV." Accessed: Oct. 01, 2023. [Online]. Available: http://www.upv.es/entidades/ASIC/seguridad/353404normalc.html

[138] "Cybersecurity Tip of the Week - Should You Click On Unsubscribe? - UTHSC News." Accessed: Sep. 30, 2023. [Online]. Available: https://news.uthsc.edu/announcements/cybersecurity-tip-of-the-week-should-you-click-on-unsubscribe/

[139] "¿Qué es una huella digital?" Accessed: Oct. 02, 2023. [Online]. Available: https://latam.kaspersky.com/resource-center/definitions/what-is-a-digital-footprint

[140] "How Do Hackers Hide?" Accessed: Oct. 06, 2023. [Online]. Available: https://www.bulletproof.co.uk/blog/how-hackers-hide

[141] "Botnet - Tech-FAQ." Accessed: Dec. 21, 2023. [Online]. Available: https://www.tech-faq.com/botnet.html

[142] "Los sistemas operativos actuales más usados y para móviles." Accessed: Nov. 30, 2023. [Online]. Available: https://elidiomadelaweb.com/los-sistemas-operativos-actuales/

[143] "¿Cuál es el sistema operativo más seguro? - Blog Prosegur." Accessed: Nov. 30, 2023. [Online]. Available: https://blog.prosegur.es/sistema-operativo-mas-seguro/

[144] "Qué es un navegador web: Características y funciones 【Guía】." Accessed: Oct. 07, 2023. [Online]. Available: https://axarnet.es/blog/navegador-web

[145] "GitHub - alrra/browser-logos: High resolution web browser logos." Accessed: Oct. 07, 2023. [Online]. Available: https://github.com/alrra/browser-logos

[146] "Los 10 mejores navegadores web de privacidad en 2023." Accessed: Oct. 07, 2023. [Online]. Available: https://es.wizcase.com/blog/mejores-navegadores-web-privacidad/

[147] "Cuál es el mejor navegador web para el móvil." Accessed: Oct. 07, 2023. [Online]. Available: https://blog.phonehouse.es/2022/03/18/mejor-navegador-movil/

[148] "Qué son las DNS, para qué sirven y por qué son tan importantes | Computer Hoy." Accessed: Dec. 01, 2023. [Online]. Available: https://computerhoy.com/reportajes/tecnologia/que-son-dns-que-sirven-que-son-tan-importantes-298499

[149] "¿Qué es y cómo funciona el servidor DNS? - Bit2Me Academy." Accessed: Dec. 01, 2023. [Online]. Available: https://academy.bit2me.com/que-es-el-servidor-dns/

[150] "Qué es un servidor DNS y porque deberías cambiarlo para proteger tu privacidad | Computer Hoy." Accessed: Dec. 01, 2023. [Online]. Available: https://computerhoy.com/ciberseguridad/servidor-dns-porque-deberias-cambiarlo-proteger-privacidad-1265140

[151] "GRC's | DNS Nameserver Performance Benchmark." Accessed: Dec. 01, 2023. [Online]. Available: https://www.grc.com/dns/benchmark.htm

[152] "DNS over TLS / HTTPS / QUIC y el futuro de la privacidad." Accessed: Dec. 01, 2023. [Online]. Available: https://ciberseguridad.blog/dns-over-tls-https-quic-y-el-futuro-de-la-privacidad/

[153] "Comprobación del navegador de Cloudflare | Cloudflare." Accessed: Dec. 01, 2023. [Online]. Available: https://www.cloudflare.com/es-es/ssl/encrypted-sni/?ref=ciberseguridad.blog#dns

[154] "What is SNI? How TLS server name indication works | Cloudflare." Accessed: Dec. 01, 2023. [Online]. Available: https://www.cloudflare.com/learning/ssl/what-is-sni/

[155] "Encrypted Client Hello - the last puzzle piece to privacy." Accessed: Dec. 01, 2023. [Online]. Available: https://blog.cloudflare.com/announcing-encrypted-client-hello/

[156] "Introducing Encrypted Client Hello (ECH)." Accessed: Dec. 01, 2023. [Online]. Available: https://www.tunnelbear.com/blog/introducing-encrypted-client-hello-ech/

[157] "Cómo evitar los bloqueos de webs de las operadoras activando ECH en Chrome." Accessed: Dec. 01, 2023. [Online]. Available: https://www.xataka.com/basics/como-evitar-bloqueos-webs-operadoras-activando-ech-chrome

[158] "5/9/14 Eyes Alliance: What You Need to Know | VeePN Blog." Accessed: Dec. 07, 2023. [Online]. Available: https://veepn.com/blog/5-9-14-eyes-alliance/

[159] "5/9/14 Eyes Countries & VPNs: What You Need to Know (2023)." Accessed: Dec. 07, 2023. [Online]. Available: https://www.vpnmentor.com/blog/understanding-five-eyes-concept/

[160] "La 'Biblia' de la Ciberseguridad." Accessed: Dec. 08, 2023. [Online]. Available: https://ciberseguridad.blog/la-biblia-de-la-ciberseguridad/#T4

[161] "Cifrado PGP, ¿Qué es y cómo funciona?" Accessed: Dec. 08, 2023. [Online]. Available: https://ciberseguridad.com/guias/prevencion-proteccion/criptografia/cifrado-pgp/

[162] "PGP: qué es y cómo funciona – Kaspersky Daily | Blog oficial de Kaspersky." Accessed: Dec. 08, 2023. [Online]. Available: https://www.kaspersky.es/blog/pgp-privacidad-seguridad-y-autenticacion-fiables-para-todos/1781/

[163] "Esteganografía. Definición, técnicas y usos frecuentes | Ayuda Ley Protección Datos." Accessed: Dec. 21, 2023. [Online]. Available: https://ayudaleyprotecciondatos.es/2021/03/17/esteganografia/#Esteganografia_en_internet_y_redes_sociales

[164] "Introducción a la Esteganografía – Proteger mi PC." Accessed: Dec. 21, 2023. [Online]. Available: https://protegermipc.net/2018/06/26/introduccion-a-la-esteganografia/

[165] "Tor vs una VPN: ¿Qué opción es más segura y privada? 2023." Accessed: Dec. 08, 2023. [Online]. Available: https://es.vpnmentor.com/blog/tor-vs-una-vpn-que-opcion-es-mas-segura-y-privada-en/

[166] "How Tor Browser Works and Where to Find Built-in Tor Bridges | Fortinet." Accessed: Dec. 08, 2023. [Online]. Available: https://www.fortinet.com/blog/threat-research/dissecting-tor-bridges-pluggable-transport

[167] "pluggable transports | Tor Project | Support." Accessed: Dec. 20, 2023. [Online]. Available: https://support.torproject.org/glossary/pluggable-transports/

[168] "Proxy Anonymity Levels - Elite, Anonymous, Transparent proxies." Accessed: Dec. 20, 2023. [Online]. Available: https://www.proxynova.com/proxy-articles/proxy-anonymity-levels-explained/

[169] "What is a digital fingerprint and how does it work?" Accessed: Nov. 09, 2023. [Online]. Available: https://smowl.net/en/blog/what-is-a-digital-fingerprint/

[170] "Cookies: Qué son, para qué sirven y tipos | AyudaLey Datos." Accessed: Nov. 09, 2023. [Online]. Available: https://ayudaleyprotecciondatos.es/cookies/

[171] "ClickDatos • ¿Qué tipos de cookies existen?" Accessed: Nov. 09, 2023. [Online]. Available: https://clickdatos.es/que-tipos-de-cookies-existen/

[172] "Qué son las 'supercookies', los nuevos sistemas de seguimiento de nuestra navegación para los que no existe un botón de borrado." Accessed: Nov. 10, 2023. [Online]. Available: https://www.genbeta.com/navegadores/que-supercookies-nuevos-sistemas-seguimiento-nuestra-navegacion-para-que-no-existe-boton-borrado

[173] "Mozilla Explains: Cookies and supercookies." Accessed: Nov. 10, 2023. [Online]. Available: https://blog.mozilla.org/en/internet-culture/mozilla-explains-cookies-and-supercookies/

[174] "Trustpid, guía a fondo: qué es, cómo funciona y cómo puedes desactivar esta publicidad de tu teléfono para que no te rastreen." Accessed: Nov. 11, 2023. [Online]. Available: https://www.xatakamovil.com/movistar/trustpid-guia-a-fondo-que-como-funciona-como-puedes-desactivar-esta-publicidad-tu-telefono-no-te-rastreen

[175] "Cookies zombie y Super cookies. Todo lo que debes saber | Ayuda Ley Protección Datos." Accessed: Nov. 10, 2023. [Online]. Available: https://ayudaleyprotecciondatos.es/cookies/zombie-supercookies/

[176] "Cómo borrar y controlar las cookies y las cookies Flash en Chrome o Firefox." Accessed: Nov. 10, 2023. [Online]. Available: https://www.genbeta.com/paso-a-paso/como-borrar-y-controlar-las-cookies-y-las-cookies-flash-en-chrome-o-firefox

[177] "GitHub - samyk/evercookie: Produces persistent, respawning 'super' cookies in a browser, abusing over a dozen techniques. Its goal is to identify users after they've removed standard cookies and other privacy data such as Flash cookies (LSOs), HTML5 storage, SilverLight storage, and others." Accessed: Nov. 10, 2023. [Online]. Available: https://github.com/samyk/evercookie

[178] "Qué pueden saber con la dirección IP y cómo ocultarla." Accessed: Nov. 11, 2023. [Online]. Available: https://www.redeszone.net/tutoriales/redes-cable/que-saben-direccion-ip/

[179] "Todo lo que pueden saber de ti gracias a la IP de tu ordenador | Lifestyle | Cinco Días." Accessed: Nov. 11, 2023. [Online]. Available: https://cincodias.elpais.com/cincodias/2015/05/25/lifestyle/1432568528_799012.html

[180] "Operadores de Búsqueda | Aprende a usar Google de verdad (2022)." Accessed: Nov. 11, 2023. [Online]. Available: https://www.thepowermba.com/es/blog/operadores-de-busqueda

[181] "Cómo encontrar a una persona por Internet: 3 métodos que funcionan." Accessed: Nov. 27, 2023. [Online]. Available: https://www.mundodeportivo.com/urbantecno/redes-sociales/como-encontrar-a-una-persona-por-internet-3-metodos-que-funcionan

[182] "15 Herramientas gratis para buscar personas en Redes Sociales." Accessed: Nov. 27, 2023. [Online]. Available: https://www.strategiaonline.es/15-herramientas-gratis-para-buscar-personas-en-redes-sociales/

[183] "Huella Digital de Internet: qué es, cómo ver y cómo buscar gratis con cualquier navegador web." Accessed: Nov. 27, 2023. [Online]. Available: https://www.xataka.com.mx/basics/huella-digital-internet-que-como-ver-como-buscar-gratis-cualquier-navegador-web

[184] "Egosurfing: ¿Qué información hay sobre mí en Internet? | Ciudadanía | INCIBE." Accessed: Nov. 27, 2023. [Online]. Available: https://www.incibe.es/ciudadania/blog/egosurfing-que-informacion-hay-sobre-mi-en-internet

[185] "Las 15 mejores alternativas a Google Alerts - Capterra España 2023." Accessed: Nov. 27, 2023. [Online]. Available: https://www.capterra.es/alternatives/202337/google-alerts

[186] "Las 20 mejores alternativas a Google Alerts - GetApp España 2023." Accessed: Nov. 27, 2023. [Online]. Available: https://www.getapp.es/alternatives/131703/google-alerts

[187] "Las 7 alternativas a Google Alerts más interesantes | Periodismo Ciudadano." Accessed: Nov. 27, 2023. [Online]. Available: https://www.periodismociudadano.com/las-7-alternativas-a-google-alerts-mas-interesantes/

[188] "Cómo limpiar tu huella digital | Avast." Accessed: Nov. 27, 2023. [Online]. Available: https://blog.avast.com/es/how-to-clean-up-your-digital-footprint-avast

[189] "QUIÉNES USAN NUESTRAS HUELLAS DIGITALES | HUELLA DIGITAL." Accessed: Dec. 06, 2023. [Online]. Available: https://moodle2021-22.ua.es/moodle/pluginfile.php/196879/mod_resource/content/12/tema/quines_usan_nuestras_huellas_digitales.html

[190] "Have I Been Pwned: Check if your email has been compromised in a data breach." Accessed: Dec. 21, 2023. [Online]. Available: https://haveibeenpwned.com/

[191] "Cómo borrar la información personal tuya que aparece en Internet." Accessed: Dec. 21, 2023. [Online]. Available: https://www.xataka.com/basics/como-borrar-informacion-personal-tuya-que-aparece-internet

[192] "Google Alerts alternative. The best alert service with Twitter results." Accessed: Dec. 22, 2023. [Online]. Available: https://www.talkwalker.com/alerts

[193] "¿De qué se habla en Internet? Mejores alternativas a Google Alerts | Computer Hoy." Accessed: Dec. 22, 2023. [Online]. Available: https://computerhoy.com/noticias/internet/que-habla-internet-mejores-alternativas-google-alerts-74801

[194] "Yahoo! Alertas es una alternativa sólida a las alertas de Google / Internet | ¡Noticias del mundo de la tecnología moderna!" Accessed: Dec. 22, 2023. [Online]. Available: https://es.ephesossoftware.com/articles/internet/yahoo-alerts-is-a-solid-google-alerts-alternative.html

[195] "Degoogling your life, a never-ending story." Accessed: Dec. 24, 2023. [Online]. Available: https://dylanvanassche.be/assets/degoogling-your-life/#acts-for-privacy-4

[196] "DeGoogling. I have been rdallman10@gmail.com since… | by Reed Allman | Medium." Accessed: Dec. 24, 2023. [Online]. Available: https://medium.com/@rdallman10/degoogling-eb3709bdfd4c

[197] "CURIA - Documentos." Accessed: Dec. 24, 2023. [Online]. Available: https://curia.europa.eu/juris/document/document.jsf?text=&docid=152065&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=269208

[198] G. Vormayr, T. Zseby, and J. Fabini, "Botnet Communication Patterns", doi: 10.1109/COMST.2017.2749442.

[199] R. Shirey, "Network Working Group".

[200] "¿Qué son los insiders y por qué son tan peligrosos para una empresa?" Accessed: Oct. 07, 2023. [Online]. Available: https://www.cybereop.com/blog/que-son-los-insiders-y-por-que-son-tan-peligrosos-para-una-empresa.html