# C1.

# INTRODUCCIÓN A LA DEEPWEB Y LA DARK WEB

## C1.1. Darknets y Deep web: diferencias

# ÍNDICE

***1*** Darknets y Deep web: diferencias

***2*** Acceso a darknets.

***3*** Búsquedas en redes alternativas / deep web

# 1 DARKNETS Y DEEP WEB: DIFERENCIAS

*DeepWeb*, *DarkNet*, *DarkWeb* … son términos que encontramos a menudo en los medios, muchas veces objeto de titulares sensacionalistas. Su contenido parece difuso y muchas veces intercambiable. Para comenzar este módulo, intentaremos clarificar esos términos.

*DeepWeb*, o web profunda es aquélla parte del web que los buscadores clásicos (*Google*, *Bing*, etc.) no son capaces de indizar. Es importante entender porqué esos buscadores no pueden indizar ciertas páginas; esto nos permitirá, en consecuencia, entender también qué es el *DeepWeb* o *Web Profunda*.

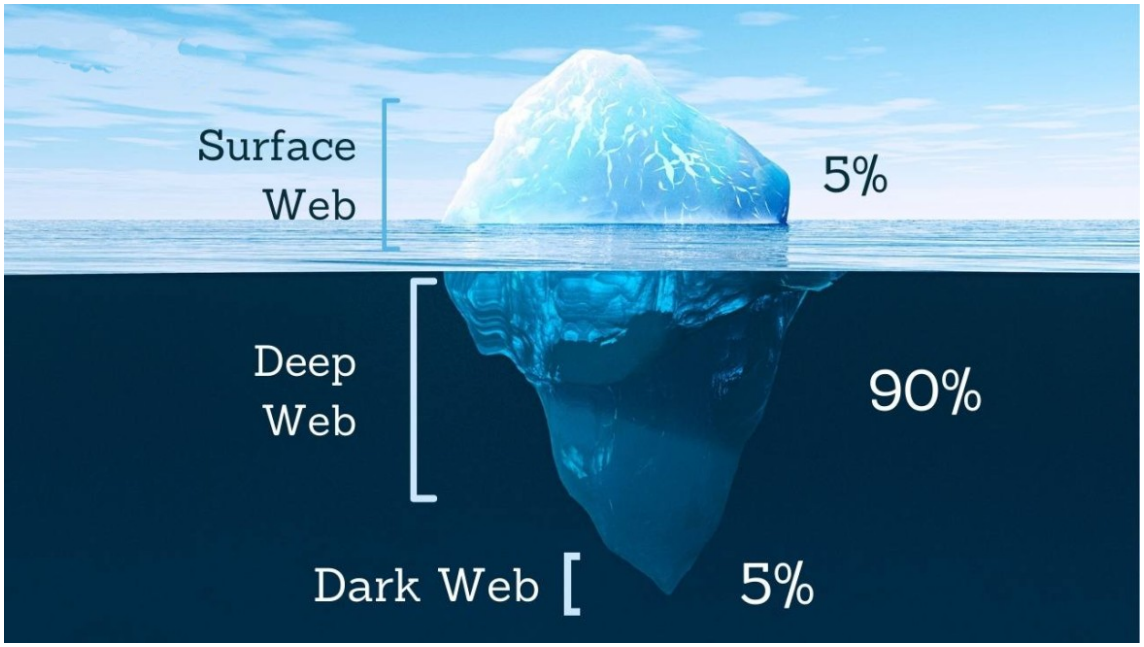
Los buscadores que utilizamos habitualmente, como Google y otros utilizan unos programas conocidos como *crawler*s ( a veces también se les llama spiders o bots, aunque estos términos tienen un significado más amplio que incluye también otro tipo de programas). Un *crawler* es un programa que recorre la red de manera autónoma y automática; en su recorrido, recolecta las páginas por las que pasa y las añade a la base de datos del buscador en cuestión. Cuando alguien hace una búsqueda en ese buscador, éste busca en su base de datos y devuelve las páginas que respondan a esa búsqueda.

Para poder efectuar ese recorrido, el *crawler* parte de una lista de direcciones web conocidas como semillas; el *crawler* toma la primera semilla de la lista, la descarga y la analiza buscando en ella enlaces a otras páginas. Los enlaces que encuentra los añade a la lista de semillas. Después, continúa con la siguiente dirección de esa lista de semillas, hace lo mismo y continúa con la siguiente semilla y así sucesivamente hasta que se acabe la lista.

Los *crawler*s sólo pueden llegar a páginas que están enlazadas desde otras, y para poder descargarlas, obtener los enlaces que contengan y añadirlas a la base de datos necesitan que esas páginas sean accesibles con sólo activar el enlace que conduce a ellas. Las intranets, por ejemplo, que requieren autenticación, no pueden ser exploradas por el *crawler*. Otras páginas no requieren propiamente autenticación, pero para acceder a ellas se requiere que el usuario realice una serie de acciones, por ejemplo cumplimentar datos de un formulario o similares.
Por supuesto, el *crawler* no podrá acceder al contenido y enlaces de páginas encriptadas o a páginas y sitios que requieren protocolos o enrutamientos diferentes de los estándares; y en algunos casos tampoco podrá acceder a algunas de lo que se conoce como páginas dinámicas.

Pues bien, toda esta parte del web, que está ahí pero que no está en las bases de datos de los buscadores es lo que llamamos *Web Profunda o Deep Web*.

Es fácil comprender que, detrás de esas páginas con autenticación, o tras un formulario de búsqueda en, por ejemplo, el catálogo de una biblioteca hay una cantidad muy importante de páginas a las que los *crawler*s de los buscadores no llegan. El tamaño de la *Web Profunda* debe ser, pues, muy grande. Haciendo extrapolaciones a partir de datos muestreados y conocidos, algunos investigadores llegan a aventurar cifras; así, el *Web Profundo* sería entre 500 y 5000 veces más grande que el *Web de Superficie*.




Como se ve, la mayor parte de ese *Web Profundo* tiene un contenido perfectamente lícito; lamentablemente, la expresión Profundo ha dado pie a excesos indocumentados que sugieren misterio, arcanos insondables y similares. No existen regiones en el *Web Profundo*, ni niveles de profundidad, ni mucho menos fosas abisales como algunas personas escriben.

## Dark Nets

Algunas partes de Internet (páginas y sitios web, pero también sistemas de mensajería, de intercambio de ficheros y otros) utilizan sistemas para proteger la privacidad de sus usuarios; por proteger la privacidad debemos entender aquí preservar el anonimato de esos usuarios. La forma básica de mantener ese anonimato es ocultar las **IP** de los usuarios de manera que no puedan ser rastreados.

Sabido es que los dispositivos a través de los cuales nos conectamos a Internet tienen un código o número **IP** que los identifica y que la información que intercambiamos viaja por la red dividida en bloques o paquetes. Cada paquete lleva, entre otras cosas, la **IP** del dispositivo origen y la **IP** del dispositivo destino; así que se puede saber qué dispositivo envía o recibe información hacia o desde qué otro dispositivo.

Muchos dispositivos tiene lo que se conoce como **IP** dinámica: el proveedor de acceso a Internet (Movistar, Jazztel, etc.) asigna una **IP** diferente a ese dispositivo cada vez que inicia una sesión de Internet. Esto permite a los proveedores dar servicio a muchos usuarios confiando en que no todos se conectarán al mismo tiempo. Pero, en cualquier caso, esos proveedores de Internet conservan registro de qué dispositivo (de qué cliente) tenía una determinada **IP** en un momento determinado.

Para ocultar la **IP** real de los usuarios se han ideado diversos sistemas que requieren el uso de un software especial. Con frecuencia ese software especial se apoya en los protocolos estándar de Internet, sobreponiéndose como una capa adicional. Las redes que aplican estos sistemas forman parte de Internet, la información que intercambian es transportada a través de Internet.

Obviamente, forman parte del *Web Profundo*. Los buscadores no pueden llegar a sus páginas y sitios; en realidad, técnicamente podrían hacerlo si ellos utilizasen también ese software especial, pero pocos lo hacen porque para acceder a las páginas encontradas es necesario también software especial. Hay alguna excepción a esto, que se verá más adelante, como el caso del buscador Ahmia.

Pues bien, el conjunto de sitios y servicios que aplican software específico para ocultar la **IP** de sus usuarios y preservar su anonimato es lo que se conoce como *Darknets*. Son varias, puesto que son varios los sistemas y los programas que se aplican para lograr esos fines.
Como se ha dicho antes, las *Darknets* o *Redes Oscuras* no solamente proporcionan páginas web, también mensajería, transferencia de archivos , etc. Como el servicio más usual, sin embargo, es el de páginas web, se habla también de *Dark Web* o *Web Oscuro*.
A menudo, *DarkNets* y *DarkWeb* se utilizan como términos intercambiables.

Las *DarkNets* forman parte de la *Web Profunda*, pero tienen un tamaño muy reducido. Debido a sus características es difícil saber su tamaño (también es difícil o imposible saber el tamaño del *Web de Superficie*); muchos de sus sitios y servicios son efímeros, otros cambian frecuentemente de dirección.

La *Red Oscura* más difundida es **Tor** (*The Onion Router*) y a ella dedicaremos las siguientes páginas, pero hay otras como *i2P*, *Freenet*, *ZeroNet*, etc.


Nuevamente, la expresión *Oscura* nos induce a prejuzgar los contenidos de estas redes. En realidad, fueron ideadas por personas deseosas de defender la privacidad personal frente al control y supervisión de gobiernos, empresas, organizaciones religiosas, etc. De hecho, la mayor parte de sus contenidos no tienen nada que ver con actividades delictivas, y carecen de interés directo por parte de las agencias para el cumplimiento de la ley (*Law Enforcement Agencies*).


Una parte de los usos y usuarios de estas redes tienen que ver con la práctica de la libertad de expresión y comunicación en países con regímenes dictatoriales que practican la censura en Internet. Otra parte, finalmente, del contenido de estas redes aprovecha el anonimato seguro para actividades ilícitas, o al menos dudosas. Algunos investigadores que han examinado exhaustivamente los

contenidos de la red **Tor** estiman que ese contenido ilícito o dudoso ocupa alrededor del 20 % de esa red.


*Ejemplo de contenido de la DarkWeb*