

# LA CARA OCULTA DE LA DARK WEB. UN ANÁLISIS VOLUMÉTRICO DE LA PRESENCIA DE MASI EN FRENET

Facundo Gallo-Serpillo<sup>1</sup>

Profesor e investigador  
Universidad Internacional de La Rioja

**Title:** *The Hidden Face of the Dark Web: A Volumetric Analysis of CSAM Presence on Freenet*

**Resumen:** *Freenet*, como parte constitutiva de la denominada *Deep Web*, conforma un ecosistema de servicios anónimos que ha sido escasamente abordado desde una perspectiva académica. Si bien existen estudios preliminares de carácter exploratorio, la mayoría de estos se limitan a la descripción telemática de la red o a la categorización general de sus servicios. En contraste, se identifican escasos antecedentes orientados específicamente a la detección y análisis de material de abuso sexual infantil (en adelante, MASI).

En este contexto, el presente estudio tiene por objetivo realizar una revisión sistemática y cuantitativa del contenido accesible en *Freenet*, con el fin de identificar servicios ocultos vinculados a la distribución de MASI. Para tal propósito, se desarrollaron e implementaron procedimientos automatizados capaces de extraer, filtrar y clasificar enlaces procedentes de los principales servicios de indexación disponibles en dicha red. Con todo ello, se pretende exponer recomendaciones prácticas para su transposición a planes y procedimientos preventivos.

---

<sup>1</sup> Doctor en criminología. Especializado en ciberseguridad, cibercrimen y ciberinvestigación. Miembro del Grupo DESONT-UNIR. facundodavid.gallo@unir.net. <https://orcid.org/0000-0002-5189-1128>.

Como resultado del análisis, se detectaron 367 sitios web relacionados con contenido de abuso sexual infantil, todos ellos localizados a través de siete indexadores activos dentro de la *Dark Web*. Estos hallazgos permiten evidenciar la persistencia de MASI en entornos anónimos y subrayan la necesidad de continuar desarrollando herramientas tecnológicas y marcos legales eficaces para su detección y persecución.

**Palabras clave:** MASI, *Dark Web*, Freenet, Cibercrimen, Anonimato.

**Abstract:** Freenet, as a constituent part of the so-called Deep Web, is an ecosystem of anonymous services that has been scarcely approached from a rigorous academic perspective. Despite the existence of preliminary exploratory studies, the majority of these are confined to the telematic description of the network or to the general categorisation of its services. Conversely, there is a paucity of evidence specifically aimed at detecting and analysing content related to child sexual abuse material (hereinafter referred to as CSAM).

In this context, the present study aims to carry out a systematic and quantitative review of the content accessible on Freenet, in order to identify hidden services linked to the distribution of child sexual abuse material. To this end, automated procedures were developed and implemented for the purpose of extracting, filtering and classifying links from the main indexing services available on the network. The aim is to provide practical recommendations for transposing them into preventive plans and procedures.

The analysis revealed 367 websites related to CSAM, all of which were located through a mere seven active indexers within the Dark Web. These findings underscore the endurance of this category of content in anonymous environments and underscore the necessity to persist in the development of technological tools and efficacious legal frameworks for its detection and prosecution.

**Keywords:** CSAM, *Dark Web*, Freenet, Cybercrime, Anonymity.

**Sumario:** Sumario. 1. Introducción. 1.1. Conceptualización *Deep* y *Dark Web*. 1.2. Freenet como *Dark Web* alternativa. 1.3. Presencia de MASI en Internet. 1.3.2. Presencia de MASI en Freenet. 1.5. Trabajos relacionados. 2. Metodología. 3. Resultados. 4. Conclusiones. 4.1. Principales resultados y discusiones. 4.2. Recomendaciones practices para una detección temprana. 5. Referencias.

## 1. Introducción

### 1.1. Conceptualización *Deep* y *Dark Web*

En los tiempos modernos, donde el acceso democrático a la información pareciera un derecho irrefutable, los usuarios hacen un uso coti-

diano de las diversas fuentes publicadas en páginas Web y redes sociales para compartir experiencias, adquirir conocimientos o intercambiar bienes y servicios; este conjunto de páginas Web accesibles desde un navegador estándar es conocido como *Surface Web*, condensando un amplio conjunto de servicios recurrentes a disposición de los usuarios consumidores. No obstante, se estima que los sitios Web a los que accedemos con normalidad representan aproximadamente el 4% de las páginas existentes en Internet, el resto del contenido disponible está oculto (Nazah *et al.*, 2020, p. 171796), y requiere de mecanismos adicionales para su acceso; este último conjunto de servicios Web se denomina *Deep Web*.

Por otro lado, entendemos como *Dark Web* al subconjunto de páginas dispuestas en la *Deep Web* cuyo principal uso está estrechamente relacionado con la publicación de servicios ilícitos (Saleem *et al.*, 2022, p. 33628); estos servicios suelen incluir venta de drogas, tráfico de armas, MASIS, robo de datos financieros, comunicaciones terroristas, etc. (Nazah *et al.*, 2020, p. 171796).

Entre los servicios tecnológicos que facilitan la disposición de contenido en la denominada *Dark Web*, la red Tor es el sistema de servidores más extendido (Bergman & Popov, 2023, p. 35914), así como también el más popular entre sus usuarios para compartir información (Nazah *et al.*, 2020, p. 171797); además, representa una solución muy robusta en términos de seguridad (Saleem *et al.*, 2022, p. 33629) en comparación con otros sistemas de *Dark Web* como pueden ser *Lokinet*, *IPFS*, *I2P* y *Freenet*.

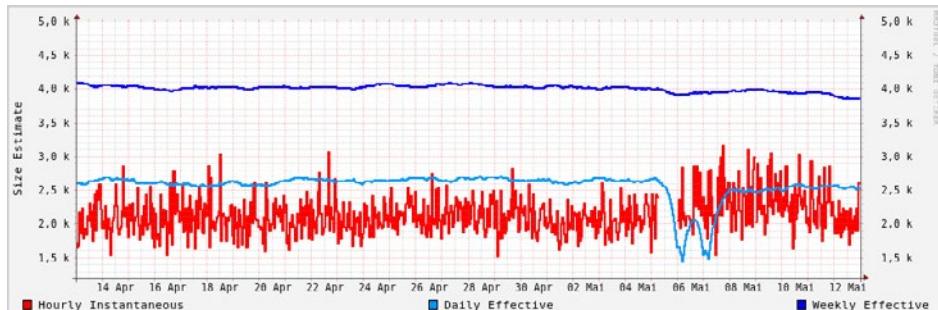
## 1.2. Freenet como Dark Web alternativa

*Freenet* es un proyecto desarrollado por Ian Clarke, y representa uno de los sistemas más extendidos de las redes anónimas, pudiendo considerarse como un repositorio de datos descentralizado para compartir información (Lee *et al.*, 2018, p. 655). Desde una perspectiva de arquitectura, la *Freenet* es una red descentralizada que permite a los usuarios compartir y acceder a información de manera anónima y resistente a la censura. Cada usuario contribuye al sistema proporcionando una parte de su ancho de banda y reservando espacio para almacenar fragmentos de datos cifrados pertenecientes a otros usuarios; este espacio se conoce como *datastore* (Figueroas-Martín *et al.*, 2022, p.4). En comparación con otras redes *peer-to-peer* (P2P), los usuarios no tienen control directo sobre el contenido almacenado en su *datastore*; en su lugar, los archivos se conservan o eliminan automáticamente en función de su popularidad, lo que contribuye a la resistencia a la censura (freenetproject, s. f.). Por otro lado, y a diferencia de lo que ocurre en Tor o I2P, en los cuales se accede al contenido Web a través de los dominios nativos *.onion* o *.i2p*, respectivamente, el acceso a los servicios Web en *Freenet* se realiza a través de

*FProxy* utilizando la siguiente nomenclatura como dirección de destino: [http://127.0.0.1:8888/<clave\\_freenet>](http://127.0.0.1:8888/<clave_freenet>); los servicios Web de *Freenet* se denominan *freesites*, y están alojados en servidores que reciben el nombre de *nodos*, los cuales representan la parte activa de la red.

Según especifican Clarke *et al.* (2001, pp. 46-66), cada *nodo* guarda fragmentos cifrados de los archivos, y el contenido se distribuye en varias localizaciones, por lo que ninguno de los contribuyentes sabe qué está guardando exactamente. Las peticiones hacia un *freesite* concreto pasan por varios *nodos*, tal como sucede en la red Tor, y cada *nodo* solo sabe del siguiente, sin tener conocimiento de quién es el origen o el destino final, por lo que resulta complejo rastrear al usuario que solicita el contenido. Finalmente, para acceder a un archivo se utiliza una clave *hash*; si se posee la clave puede encontrarse el archivo, aunque se desconozca quién lo subió y en qué nodo se ubica.

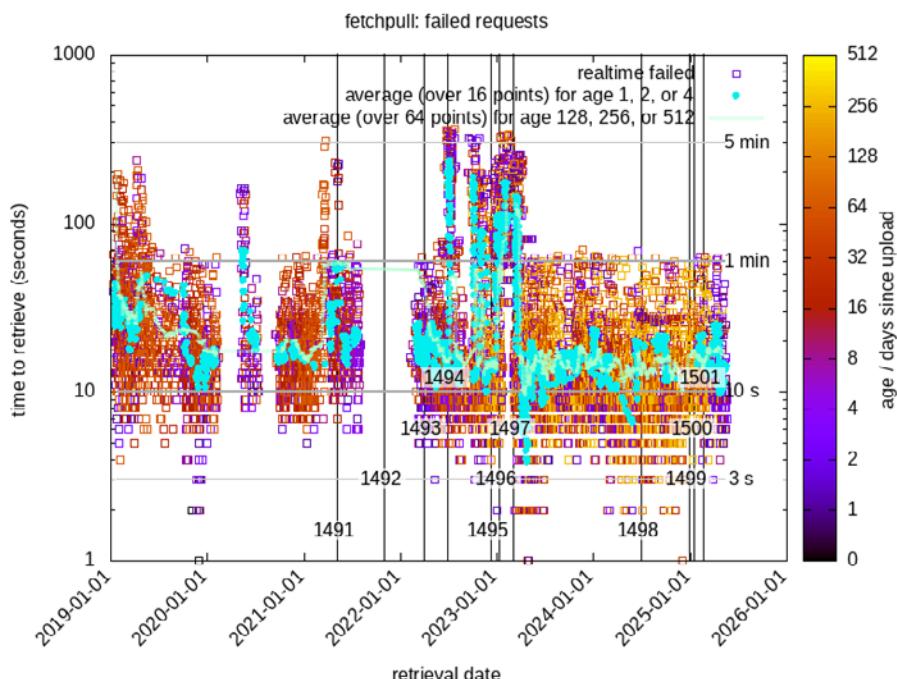
Dada la naturaleza de su diseño, el tamaño de la red *Freenet* sigue representando un dato complejo de obtener, no obstante, puede acudirse a los servicios estadísticos de la propia red que buscan representar el consumo basado en la inspección telemática; una de estas fuentes permite realizar una estimación en tiempo real sobre la cantidad de *nodos* que se encuentran disponibles en el momento de la consulta (ver Figura 1) suponiendo, eso sí, que los resultados provienen siempre de nodos seleccionados aleatoriamente en la red.



**Figura 1.** Tamaño estimado de la red Freenet (intervalo de un mes)

Fuente: <http://localhost:8888/freenet:USK@WMa1Z40iYdZZ51yctQ3toFl9zuuFEnNdsrnNejJU5KE,jCBCaNBKD5~sSQeSkyKz737Bh5ibBGqdzfD8mgfdMY,AQACAAE/statistics/802/>

No solo el volumen de nodos activos es clave para la difusión de contenido en *Freenet*, sino también el tiempo de disponibilidad de los ficheros compartidos. El siguiente gráfico, basado en un servicio estadístico de la *Freenet*, representa la prevalencia de ficheros a lo largo del tiempo; en este sentido, y al comparar el color en la línea de 128 días (por encima de 4 meses), con el color en la línea de 1 día, puede apreciarse cuántas inserciones siguen activas después de 128 días.



**Figura 2.** Histograma de disponibilidad de ficheros en Freenet

Fuente: <http://localhost:8888/freenet:USK@lwR9sLnZD3QHveZa1FB0dAHgeck~dFNBg368mY09wSU,0Vq~4FXSUjI-op3QdzqjZsIvrNMYWlnSdUwCl-ZIfYA,AQACAAE/fetchpullstats/1244/>

### 1.3. Presencia de MASI en Internet

#### 1.3.1. Contexto y recorrido histórico

El contenido relativo a MASI continúa su proliferación en Internet, victimizando continuamente a los menores explotados; en esta dirección, Najat M'jid Maala, relatora de la ONU, afirma que «*unos 750 mil pedófilos en el mundo están conectados permanentemente a la red y que existen 4 millones de sitios web con contenidos pornográficos que exponen niños*»; además, la venta de MASI representa un lucrativo negocio que mueve en torno a 20.000 millones de dólares por año (Parra González, 2016, p. 28). Dada la innegable criticidad de su naturaleza del delito (al verse involucrado un colectivo vulnerable), detectar servicios de MASI se ha convertido en una de las prioridades estratégicas de las fuerzas y cuerpos de seguridad.

Desde mediados del siglo XX hasta la entrada del siglo XXI, la proliferación de MASI en Internet se ha visto propiciada por la facilidad de acceso y creación de páginas Web, erigiéndose como principal medio de difusión por encima de otros mecanismos tecnológicos (García Rojo, 2001, p. 1). Esto se debe a que el alojamiento Web proporciona una forma ágil de visibilizar y difundir contenido a un bajo coste, democratizando su uso para diversas finalidades; en consecuencia, puede afirmarse que la disposición de páginas Web se corresponde con la primera fase en la difusión de MASI en Internet (García Rojo, 2019, p. 222).

Con la entrada del siglo XXI, las redes P2P se consolidaron como una vía alternativa para la distribución de contenido gratuito online, gracias a que su arquitectura permite operar bajo el amparo de una red de servicios distribuidos y descentralizados (Wolak *et al.*, 2014, p. 348), características intrínsecas que permiten diluir la responsabilidad penal entre una gran cantidad de usuarios que operan desde diversas localizaciones geográficas. En la actualidad, las redes P2P comparten protagonismo junto a la red Tor debido, una vez más, a su uso público y gratuito, así como también a las facilidades que brindan para la comisión del delito (anonimato, bajo coste de publicación Web, pago con criptodivisas, etc.); en esta dirección, y conforme a lo reseñado en el informe de la Fundación Alia2, España representa el segundo país con mayor implicación en el tráfico de MASI dentro de las redes P2P (Coto & França Tarragó, 2014, p. 56). Además, Europol, en su informe IOCTA (Europol, 2021, p. 25), también pone el foco en las redes P2P como uno de los principales mecanismos de distribución.

Finalmente, como principal desafío contra la prevención del MASI, Europol señala la introducción de IA para la generación de MASI simulado. La creación de este tipo de material plantea desafíos significativos, particularmente en lo relativo a la identificación de víctimas reales y a la determinación del marco jurídico aplicable a la investigación. Aun en aquellos supuestos en los que el contenido es íntegramente artificial y no involucra a menores reales, el material generado mediante inteligencia artificial sigue contribuyendo a la cosificación y sexualización de la infancia; además, la proliferación de este contenido simulado incrementa el volumen de MASI en circulación, lo que dificulta de manera sustancial tanto la identificación de posibles víctimas como la atribución de responsabilidad penal a los autores (Europol, 2024, p. 24).

### 1.3.2. Presencia de MASI en Freenet

En lo que respecta a la *Dark Web*, las páginas relacionadas con distribución de MASI conforman los servicios más consumidos por sus usuarios (Owen y Savage, 2015); más en concreto, el contenido relacionado con MASI en la *Dark Web* ha sido objeto de interés en estudios recientes, aunque gran parte de ellos se circunscriben a la denominada red Tor

(Gannon *et al.*, 2023; Koebe *et al.*, 2024), y si bien es *vox populi* la presencia de MASI en *Freenet*, gran parte de su comunidad no apoya la distribución de este contenido u otras actividades ilegales, priorizando el uso del este sistema de red anónimo para la libertad de opinión (United States Sentencing Commission, 2012, p. 60).

En cuanto a los mecanismos que posibilitan la presencia de MASI en la *Freenet*, existen sistemas conocidos como *nodos índice* o *nodos de indexación* que son importantes dentro de la estructura de conectividad de la red ya que actúan como una especie de *hub* de conexión entre muchas páginas Web (Figueras-Martín *et al.*, 2022, p. 15), incluidas aquellas que disponen de material ilícito; en este sentido, y como bien se ilustra en la Figura 3, existe un conjunto de *nodos indexadores* que recopilan material relacionado con MASI, entre ellos *Child Models* o *Porndexter*, localizados durante la elaboración del presente estudio.

The screenshot shows a web page titled "Child Models Activelink Index". At the top, there is a navigation bar with links: "Activelinks <1", "TPJ <1", "TMV & Silver Sites <1", "Other <1", and "Change Log". Below the navigation bar, there are several sections listing child modeling agencies:

- A Little Agency**: Shows a small profile picture of a child's face.
- Dolce Modz**: Shows a logo for "BABY Models" and a link to "Dolce Star - Videos".
- Dream Studio**: Shows a logo for "Models Around".
- ART-Modeling Studios**: Shows logos for "Bianka", "CHERISH CHERISH", "Hermione Model fro", "Liliane", "LOLLYMODEL", "Peach Model", and "Vegas".
- Newstar & Tinymodel**: Shows logos for "HS-Alexandr", "NewStar-Kitty", "Natalie", "sweet", and "TM-Sonny".
- PR Models**: Shows a small profile picture of a child's face.

**Figura 3.** Página Child Models Activelink Index

Fuente: Información obtenida de la página *Child Models*.

En el siguiente apartado se enumera un conjunto de trabajos previos que buscan evaluar la complejidad de la red *Freenet* y brindar una aproximación volumétrica a los tipos de servicios disponibles, estén o no relacionados con la distribución de MASI.

#### 1.4. Trabajos relacionados

En lo que respecta a investigación empírica realizada en la *Freenet* debe subrayarse la poca literatura académica existente acerca de su estructura y clasificación de contenido; a modo ilustrativo, bajo una búsqueda realizada en Scopus<sup>2</sup>, mediante el empleo del término *Freenet*, y limitando los resultados a áreas específicas de computación, ingeniería y ciencias sociales, se han hallado tan solo 23 artículos publicados en los últimos 15 años. Entre los artículos recopilados, y desde una perspectiva de comunicaciones, puede hallarse literatura limitada -pero suficiente- para comprender la arquitectura establecida en *Freenet*; en este sentido, existe un estudio que verifica la capacidad de desplegar nodos en el circuito para manipular e inspeccionar la interacción de los usuarios (Xu *et al.*, 2024, p1) facilitando mecanismos de desanonimización, siendo este último tópico ampliamente abordado en estudios como el conducido por Lu *et al.* (2019). Por otro lado, cabe destacar una investigación empírica que desvela el uso de *Machine Learning* como estrategia para clasificar el tráfico y así identificar hosts dentro del circuito, aportando una visión volumétrica aproximada de los dispositivos que conforman la arquitectura de comunicaciones en *Freenet* (Lee *et al.*, 2018, p. 656). Con independencia del invaluable conocimiento que ha legado este conjunto de investigaciones, resulta evidente que ninguno arroja datos relativos a la difusión de MASI en nodos de *Freenet*.

Ahora bien, desde la perspectiva de clasificación de contenido, un antecedente directamente relacionado con el presente estudio es el llevado a cabo por Figueras-Martín *et al.* (2022) mediante el cual se ha podido demostrar la disposición de MASI en *Freenet*, haciendo constar un total 11,95% sobre el contenido inspeccionado (Figueras-Martín *et al.*, 2022, p. 11) aunque, por otra parte, no se cuantifica el contenido relacionado con MASI; además, no se especifica método alguno para la correspondiente supresión de falsos positivos, esto es, dentro de la búsqueda mediante *keywords* pueden existir coincidencias sacadas de contexto que requieren de una depuración posterior. Por ejemplo, la obtención de resultados con el término CP (abreviatura de *Child Porn*), y al no descartar sus negociaciones (no CP) contabilizar erróneamente un *match*.

Con todo ello, se ha evidenciado la necesidad de abordar el presente estudio con el fin de realizar una evaluación volumétrica orientada a la presencia de MASI en *Freenet*, dando continuidad a los antecedentes descritos y otorgando así una lectura específica ante esta tipología de servicios ilícitos.

---

<sup>2</sup> <https://www.scopus.com/>

## 2. Metodología

Se ha realizado una revisión sistemática y automatizada de 6 *nodos de indexación* presentes en *Freenet*, incluidos *nodos* destinados íntegramente a la enumeración de servicios de MASI. Los nodos seleccionados se corresponden a los principales portales de indexación, los cuales condensan y categorizan páginas Web que se encuentran actualmente disponibles.

En esta dirección, se ha procedido a crear un conjunto de *Crawlers*<sup>3</sup>, correspondiéndose estos a automatismos diseñados expresamente para el rastreo y la extracción de datos presentes en páginas Web; el desarrollo de *Crawlers* responde a la aplicación técnica de *Web Scraping* (Bergman & Popov, 2023) utilizando *beautiful soup* como principal librería de código para realizar la inspección de etiquetas en estructuras HTML que, por lo general, contienen titulares descriptivos relacionados con enlaces hacia otros *freesites*. La lectura de elementos Web en los nodos de indexación ha permitido identificar datos relevantes de manera automática, así como también filtrar la extracción de contenido en base a un array de palabras clave [*child pornography*, *child porn*, *cp*, *loli*, *lola*, *lolli*, *childabuse*, *child abuse*, *child models*, *preteen*] (ver Figura 4). Adicionalmente, se ha procedido a almacenar el total de entradas coincidentes para el elemento Web inspeccionado obteniendo, de esta manera, una aproximación volumétrica de total del contenido indexado en los diversos nodos.

```

import requests
from bs4 import BeautifulSoup
import re
import csv
from datetime import datetime

# Simulación del uso de una librería 'output'
class Output:
    @staticmethod
    def log(msg):
        print(f"[LOG] {msg}")

# Configuración
URL = "http://localhost:8888/freenet:USK@Isel-izgllcBsr~1reXUJZ"
KEYWORDS = ["child pornography",
            "child porn",
            "cp",
            "loli",
            "lola",
            "lolli",
            "childabuse",
            "child abuse",
            "child models",
            "preteen"] # palabras clave

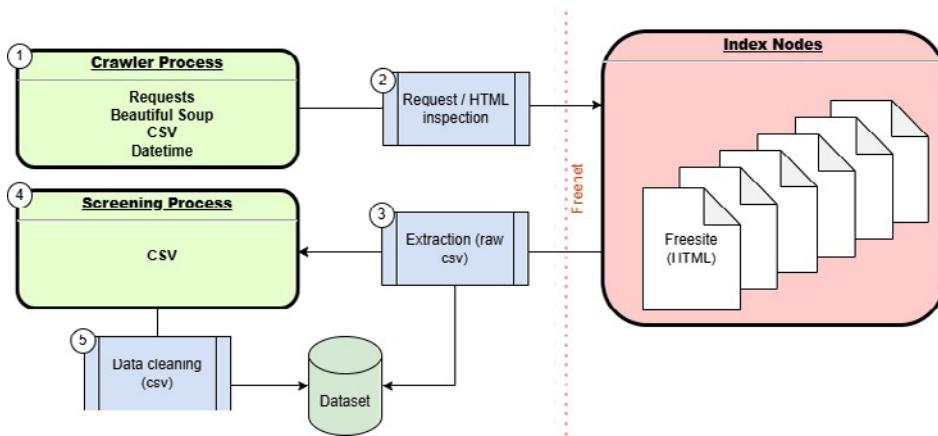
```

**Figura 4.** Crawler diseñado para Freenet

Fuente: Elaboración propia.

<sup>3</sup> <https://gitlab.com/fdgallo/freenet-web-crawlers>

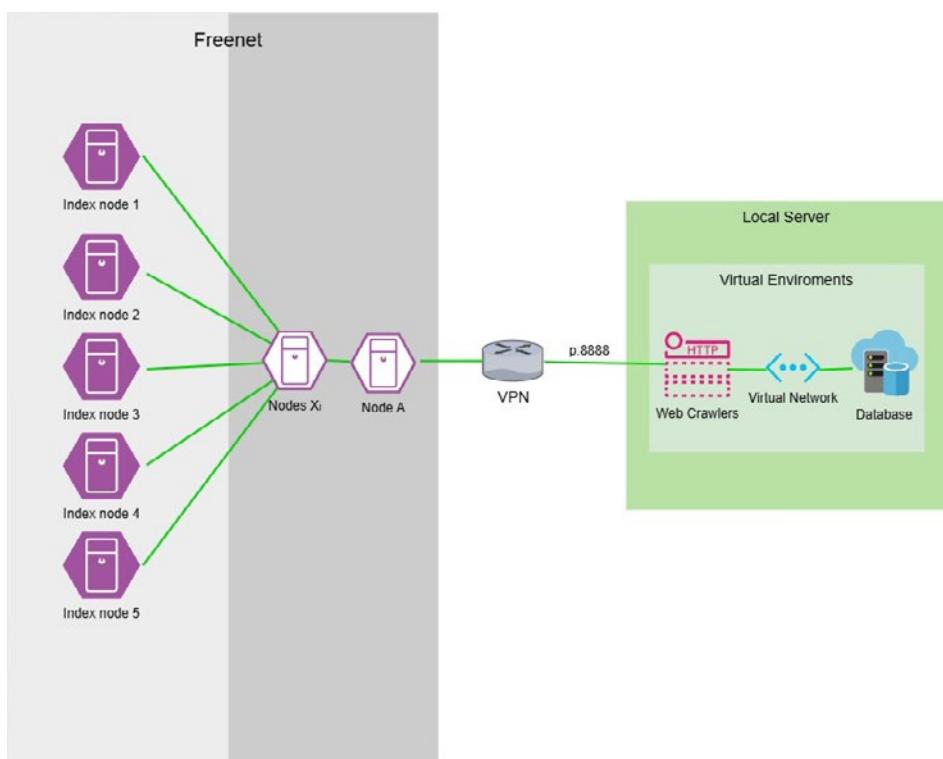
Finalmente, se ha procedido a suprimir falsos positivos mediante un cribado de datos (*data screening*), lo cual ha permitido suprimir aquellas entradas que, bajo una lectura más amplia del contexto semántico, y aun siendo datos coincidentes con una o más palabras clave, pueden ser descartados al no establecerse relación con MASI; para ello, se ha aplicado una limpieza automática de los *datasets* mediante un script de filtrado adicional que contempla los términos clave: [*not cp, no cp, not child pornography, no child pornography, not child porn, no child porn, not childabuse, no childabuse, not child abuse, no child abuse*]. A continuación, se ilustra el flujo de ejecución completo a alto nivel.



**Figura 5.** Flujograma de ejecución

Fuente: Elaboración propia.

A nivel de infraestructura, el conjunto de *Crawlers* fueron ejecutados por lote desde un entorno virtualizado con acceso a la *Freenet* mediante el puerto 8888 y un servicio VPN para proteger la identidad del investigador; los *datasets* obtenidos fueron almacenados en una base de datos no relacional.



**Figura 6.** Despliegue de servicios

Fuente: Elaboración propia.

### 3. Resultados

La ejecución de los *Crawlers* ha permitido inspeccionar de manera automatizada  $n = 46.003$  entradas de contenido en los nodos de indexación, esto es, títulos descriptivos relacionados con enlaces hacia otros *freesites*; del total de entradas extraídas (*raw*),  $n = 379$  guardan relación con los términos de búsqueda; finalmente, y tras un proceso de supresión de falsos positivos (*clean*), se han contabilizado un total de  $n = 367$  etiquetas con referencia directa a contenido MASI.

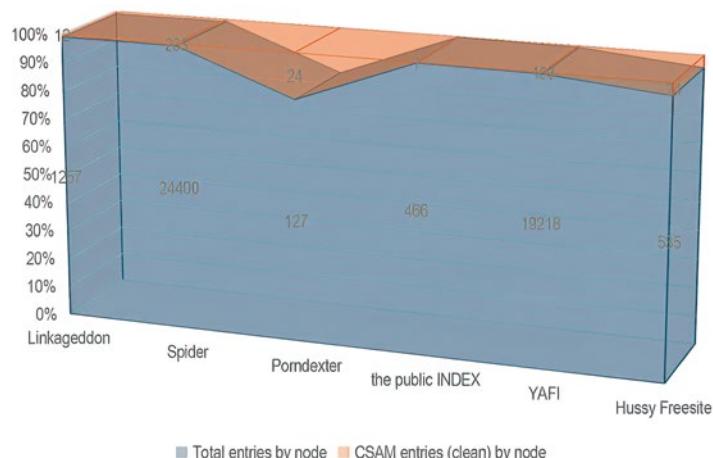
A continuación, se ofrece una tabla con la relación entre los nodos indexadores consultados, la cantidad de elementos extraídos (*raw*, *clean*), y las coincidencias en cada indexador en términos porcentuales.

**Tabla 1. Datos obtenidos de los nodos de indexación**

	Index nodes					
	<i>Linkageddon</i>	<i>Spider</i>	<i>Porndexter</i>	<i>the public INDEX</i>	<i>YAFI</i>	<i>Hussy Freesite</i>
Total entries by node	1257	24400	127	466	19218	535
CSAM entries (raw) by node	15	208	24	1	107	24
CSAM entries (clean) by node	12	203	24	1	103	24
% CSAM entries by node	0.954 %	0.008 %	0.188 %	0.002 %	0.005 %	0.044 %

Fuente: Elaboración propia.

Finalmente, y bajo una lectura pormenorizada, podemos apreciar que solo existe en un pequeño conjunto de indexadores cuya presencia de MASI es  $\geq 1\%$ , siendo *Porndexter* el servicio más notorio con  $n = 127$  entradas y  $n = 24$  enlaces Web relacionados estrechamente con MASI (ver Figura 7). Con todo ello, y en cómputos globales, se confirma que la presencia de MASI representa el 0,8% del contenido inspeccionado en los principales nodos de indexación de la *Freenet*. Estos resultados no hacen más que confirmar la prevalencia de MASI en la *Freenet* desde una aproximación empírica.

**Figura 7. Presencia de MASI en los nodos indexadores**

Fuente: Elaboración propia.

## 4. Conclusiones

### 4.1. Principales resultados y discusiones

Mediante la presente investigación se ha podido evidenciar de manera inequívoca la existencia de mercados negros orientados a la venta y distribución de MASIs en *Freenet*, identificando un total de 0,8 % de contenido relacionado con MASIs sobre el total de páginas inspeccionadas; asimismo, se ha demostrado una vez más el uso potencial de *Crawlers* para la extracción de datos en entornos heterogéneos cuando no existen otros mecanismos que faciliten la recopilación. El conjunto de servicios detectados no deja de suponer un reto constante para las fuerzas y cuerpos de seguridad en materia de prevención, detección temprana y reacción, debido, entre otros factores, a las capacidades de la *Freenet* para ocultar contenido y salvaguardar la identidad de los usuarios que consumen o distribuyen MASIs desde su red de *nodos*.

Este estudio constituye una base firme para futuras investigaciones y desarrollos, siendo una de las potenciales líneas de trabajo la combinación de *Crawlers* con modelos de *Machine Learning* con el objetivo de detectar y clasificar automáticamente el contenido publicado en los *nodos de indexación* en tiempo real; esta integración permitiría construir una herramienta con alto potencial operativo, capaz de procesar grandes volúmenes de datos de forma eficiente y con precisión. De esta manera, se facilitaría la identificación de patrones, tendencias y posibles anomalías en el contenido indexado, contribuyendo significativamente al avance de soluciones tecnológicas inteligentes aplicables en múltiples contextos desde el ámbito del cibercrimen.

### 4.2. Recomendaciones prácticas para una detección temprana

Como bien se ha hecho constar a lo largo del artículo, la proliferación de contenidos ilícitos en Internet, especialmente en la denominada *Dark Web*, constituye uno de los principales retos para la criminología contemporánea, así como también para las fuerzas y cuerpos de seguridad. La naturaleza anónima y descentralizada de estas plataformas dificulta la detección temprana del material ilícito, favoreciendo la persistencia y expansión de fenómenos como la distribución de MASIs. En este contexto, la aplicación de análisis automatizado mediante *Web Scraping* constituye una decisión necesaria y estratégica a nivel preventivo.

Desde una perspectiva criminológica, el despliegue de *Crawlers* permite no solo la identificación proactiva de MASIs, sino también el estudio de patrones de comportamiento delictivo y la anticipación de tendencias emergentes. Para las fuerzas del orden, su integración en planes de prevención supone pasar de una respuesta reactiva centrada en la investiga-

ción, hacia una intervención temprana basada en inteligencia criminal. Esta aproximación preventiva contribuye a reducir los riesgos de victimización, optimizar recursos operativos y reforzar la cooperación internacional frente a delitos transnacionales altamente dinámicos. A continuación, se enumeran dos propuestas relativas a la integración del *Web Scraping* dentro un marco preventivo frente a la distribución de MASI en *Freenet*.

– Sistema de monitorización preventivo

La presente metodología puede ser trasladada al desarrollo y despliegue de herramientas automatizadas de *Web Scraping*, idealmente en modo de microservicios capaces de rastrear foros y *marketplaces* en *Freenet* donde circula contenido MASI; estas herramientas deberán configurarse con patrones lingüísticos, metadatos técnicos y huellas digitales hashes previamente identificados por unidades especializadas.

Desde una perspectiva criminológica, este enfoque posibilita detectar tendencias de consumo y distribución de MASI, facilitando posteriormente el *crime mapping* en torno a redes y actores implicados en la difusión masiva.

– Adopción como metodología estratégica en protocolos de inteligencia criminal y prevención

Por otra parte, resulta imperativa la incorporación de los resultados del *Web Scraping* en sistemas de inteligencia criminal, o lo que es lo mismo, cruzar y contrastar los datos obtenidos contra las bases de datos de víctimas y sospechosos conocidos, así como con informes internacionales (Europol, Interpol, etc.).

En síntesis, mientras que el análisis criminológico de la información permite identificar patrones de riesgo (nuevas plataformas de distribución, métodos de anonimización emergentes, comunidades, etc.), en lo que a fuerzas y cuerpos de seguridad se refiere, la integración y correlación de datos introduce mejoras sustanciales en los protocolos operativos, potenciando la detección temprana y automatizada frente a estrategias reactivas.

## 5. Referencias

- BERGMAN, J., & POPOV, O. B. (2023). Exploring Dark Web Crawlers: A Systematic Literature Review of Dark Web Crawlers and Their Implementation. *IEEE Access*, 11, 35914-35933. <https://doi.org/10.1109/ACCESS.2023.3255165>

- CLARKE, I., SANDBERG, O., WILEY, B., & HONG, T. W. (2001). Freenet: A distributed anonymous information storage and retrieval system. pp 46-66.
- COTO, S. D., & FRANÇA TARRAGÓ, O. (2014). Flujo de material pornográfico infantil online. Estudio exploratorio en 10 países de América Latina con foco en Uruguay. Universidad de La Rioja, 1, 55–67. <https://dialnet.unirioja.es/servlet/articulo?codigo=5052359>
- EUROPOL. (2021). Internet Organised Crime Threat Assessment (IOCTA) 2021. Europol. <https://doi.org/10.2813/113799>
- EUROPOL. (2024). Internet Organised Crime Threat Assessment (IOCTA) 2024. Europol. <https://doi.org/10.2813/442713>
- FIGUERAS-MARTÍN, E., MAGÁN-CARRIÓN, R., & BOUBETA-PUIG, J. (2022). Drawing the web structure and content analysis beyond the Tor darknet: Freenet as a case of study. *Journal of Information Security and Applications*, 68, 103229. <https://doi.org/10.1016/j.jisa.2022.103229>
- FREENETPROJECT. (s. f.). Documentation [Freenetproject]. Recuperado 27 de mayo de 2025, de <https://staging.freenetproject.org/es/pages/documentation.html?>
- GANNON, C., BLOKLAND, A. A. J., HUIKURI, S., BABCHISHIN, K. M., & LEHMANN, R. J. B. (2023). Child sexual abuse material on the darknet. *Forensische Psychiatrie, Psychologie, Kriminologie*, 17(4), 353-365. <https://doi.org/10.1007/s11757-023-00790-8>
- GARCÍA ROJO, J. C. (2001). Pornografía infantil en internet. *Boletín Criminológico*, 37(3), 217–223
- GARCÍA ROJO, J. C. (2019). La realidad de la pornografía infantil en internet. *Revista de Derecho Penal y Criminología*, (9), 211–251. Recuperado a partir de <https://revistas.uned.es/index.php/RDPC/article/view/24805>
- KOEBE, T., DEL VILLAR, Z., NUTAKKI, B., SAGIMBAYEVA, N., & WEBER, I. (2024). Unveiling local patterns of child pornography consumption in France using Tor. *Humanities and Social Sciences Communications*, 11(1), 807. <https://doi.org/10.1057/s41599-024-03343-4>
- LEE, S., SHIN, S., & ROH, B. (2018). Classification of Freenet Traffic Flow Based on Machine Learning. *Journal of Communications*, 654-660. <https://doi.org/10.12720/jcm.13.11.654-660>
- LU, T., DU, Z., & JANE WANG, Z. (2019). A Survey on Measuring Anonymity in Anonymous Communication Systems. *IEEE Access*, 7, 70584-70609. <https://doi.org/10.1109/ACCESS.2019.2919322>
- NAZAH, S., HUDA, S., ABAWAJY, J., & HASSAN, M. M. (2020). Evolution of Dark Web Threat Analysis and Detection: A Systematic Approach. *IEEE Access*, 8, 171796-171819. <https://doi.org/10.1109/ACCESS.2020.3024198>

- OWEN, G., & SAVAGE, N. (2015). CIGI. Recuperado el 22 de 4 de 2024, de [https://www.cigionline.org/sites/default/files/no20\\_0.pdf](https://www.cigionline.org/sites/default/files/no20_0.pdf)
- PARRA GONZÁLEZ, A. (2016). Pornografía Infantil. Contexto Socio/Criminológico y Jurídico. Interacción y Perspectiva: Revista de Trabajo Social, 6(1), 23–41
- SALEEM, J., ISLAM, R., & KABIR, M. A. (2022). The Anonymity of the Dark Web: A Survey. IEEE Access, 10, 33628-33660. <https://doi.org/10.1109/ACCESS.2022.3161547>
- UNITED STATES SENTENCING COMMISSION. (2012). Report to the Congress: Federal Child Pornography Offenses. <https://www.ussc.gov/research/congressional-reports/2012-report-congress-federal-child-pornography-offenses>
- WOLAK, J., LIBERATORE, M., & LEVINE, B. N. (2014). Measuring a year of child pornography trafficking by U.S. computers on a peer-to-peer network. Child Abuse and Neglect, 38(2), 347–356. <https://doi.org/10.1016/j.chabu.2013.10.018>
- XU, Y., YANG, M., LING, Z., LIU, Z., GU, X., & LUO, L. (2024). A De-anonymization Attack against Downloaders in Freenet. IEEE INFOCOM 2024 - IEEE Conference on Computer Communications, 1-10. <https://doi.org/10.1109/INFOCOM52122.2024.10621209>