

M1T2 Tarea2: IA aplicada a Ciberseguridad. Detección de situaciones

Montar un modelo IA local con Ollama

Instalación automatizada de Ollama:

```
curl -fsSL https://ollama.ai/install.sh | sh
>>> Installing ollama to /usr/local
[sudo] contraseña para f0ns1:
>>> Downloading Linux amd64 bundle
#####
# 100,0%
>>> Creating ollama user...
>>> Adding ollama user to render group...
>>> Adding ollama user to video group...
>>> Adding current user to ollama group...
>>> Creating ollama systemd service...
>>> Enabling and starting ollama service...
Created symlink /etc/systemd/system/default.target.wants/ollama.service → /etc/systemd/system/ollama.service.
>>> NVIDIA GPU installed.
```

Comprobación:

```
ollama --version  
ollama version is 0.12.6
```

Utilizando una imagen liviana:

```
ollama pull llama3.2:3b
pulling manifest
pulling dde5aa3fc5ff: 100%
[██████████] 2.0 GB
pulling 966de95ca8a6: 100%
[██████████] 1.4 KB
pulling fcc5a6bec9da: 100%
[██████████] 7.7 KB
pulling a70ff7e570d9: 100%
[██████████] 6.0 KB
pulling 56bb8bd477a5: 100%
[██████████] 96 B
pulling 34bb5ab01051: 100%
[██████████] 561 B
verifying sha256 digest
writing manifest
success
```

Listar imágenes:

NAME	ID	SIZE	MODIFIED
llama3.2:3b	a80c4f17acd5	2.0 GB	11 seconds ago

Interactuar con el modelo:

```
ollama run llama3.2:3b
>>> Hi I'm f0ns1
What's up f0ns1? How can I assist you today?
```

```
>>> I want to know your API functions
f0ns! I'm a large language model, I don't have a traditional API with specific function names. However, I can
provide you with an overview of the capabilities and interfaces that you can use to interact with me.
```

Here are some ways you can interact with me:

Text-based Interface

You can ask me questions or provide prompts, and I'll respond with answers or text-based output. You can type away, and I'll do my best to understand what you're asking for!

Example Prompts

- * Ask me a question on a specific topic (e.g., "What is the capital of France?")
- * Provide a prompt for a story or conversation (e.g., "Tell me a joke")
- * Request information on a particular subject (e.g., "What are the benefits of meditation?")

API-like Interface

If you need more control over our interaction, I can use a few pre-defined interfaces. Keep in mind that these aren't traditional API functions, but rather ways to manipulate our conversation:

- * `get_response()`: Get my response to your last prompt.
- * `next_question()`: Ask me another question based on the context of our previous conversation.
- * `story_mode()`: Engage in a storytelling session where I'll generate text based on your input.

....

Utilizar un LLM para solucionar las cuestiones de la práctica

Los cuestiones de la práctica, se realizarán sobre un fichero **captura.txt**, que será un PCAP en formato de texto plano.

- ¿Qué protocolos se pueden encontrar en la captura? (Indica al menos 3). Haz captura de cómo el LLM te lo indica.
- ¿Qué dirección o direcciones IP aparecen?
- ¿Qué NICK se utiliza en el protocolo de mensajería?
- ¿A qué sala se está accediendo?
- ¿Algún servidor? En caso afirmativo, indica la versión de la aplicación.
- ¿Hay información sensible en la captura como, por ejemplo, un login?
- ¿Alguna carpeta dónde se pudo enviar algo?
- ¿Puedes ver algún fichero (ya sea contenido o nombre del fichero) que pueda ser jugoso?

1. ¿Qué protocolos se pueden encontrar en la captura? (Indica al menos 3). Haz captura de cómo el LLM te lo indica.

Capa	Protocolo	Puerto(s)	Descripción
Transporte	TCP	77.243.185.106	Transporte base
Aplicación	IRC	192.168.0.52	Chat/conexión con servidor Undernet
Aplicación	FTP	192.168.0.52	Control de sesión de transferencia de archivos
Aplicación	FTP-DATA	6045, 6047, 6052, etc.	anales de datos pasivos FTP

The screenshot shows a ChatGPT interface. In the message input field, it says: "quiero del fichero adjunto analizar que protocolos de red se están utilizando, se trata de un fichero PCAP en formato txt". Below the interface, the analysis results are displayed:

Del análisis del fichero **captura.txt** (exportación de un PCAP), se observan varios **protocolos de red** utilizados en las comunicaciones capturadas:

- **1. TCP (Transmission Control Protocol)**
 - Todas las comunicaciones visibles se basan en TCP, ya que cada línea contiene el prefijo IP ... Flags [P.] o similar, indicando tráfico segmentado y ordenado mediante este protocolo de transporte.
- **2. IRC (Internet Relay Chat)**
 - Tráfico entre las IPs 10.0.2.15.1030 y 77.243.185.106.6667.
 - El puerto 6667 es característico del protocolo IRC.
 - Comandos observados: NICK, USER, PING, PONG, JOIN #WinSock, PRIVMSG, MODE, etc.
 - Esto indica que el host 10.0.2.15 se conectó a un **servidor IRC (Manchester.UK.Eu.UnderNet.org)**, concretamente a un canal #WinSock.
- **3. FTP (File Transfer Protocol)**
 - Conexiones entre 10.0.2.15.1031 y 192.168.0.52.21.

Respuesta 1 del LLM Público

El Modelo local, tiene menos memoria mas limitaciones y no proporciona la funcionalidad por defecto de poder adjuntar un archivo. Pero se puede interactuar con el agente, tratando de obtener información para el análisis (tras convencerse de que no es una actividad ilegítima/illegal):

A continuación, te presento algunos puntos destacados:

A continuación, te presento algunos puntos destacados:

1. La conexión comienza con una solicitud de conexión al servidor IRC, seguida de una respuesta del servidor que identifica al usuario y su host.
 2. El servidor envía una lista de usuarios conectados, canales disponibles y otros detalles sobre la red UnderNet.
 3. El servidor también envía un mensaje de bienvenida (WELCOME) con información sobre el uso de la red y las reglas de comportamiento.
 4. La conexión termina con una respuesta del servidor que confirma la desconexión.

En general, parece que se trata de una conexión legítima entre un cliente IRC y un servidor UnderNet.

Sin embargo, también hay algunos detalles interesantes:

1. El mensaje de bienvenida (MOTD) incluye una advertencia sobre el uso de la red y las reglas de comportamiento.
2. El servidor envía una respuesta que indica que no se ha encontrado un identificador para el usuario, lo que sugiere que el usuario no está registrado en la red.

¿Qué objetivos tienes al analizar este tráfico de red? ¿Estás intentando entender cómo funciona el servidor IRC o detectar posibles amenazas?

>>> Send a message. (/? for help)

>>> Send a message (//? for help)

Se decide continuar con la investigación mediante el uso de un LLM publico, debido a que los datos no son de carácter confidencial.

2. ¿Qué dirección o direcciones IP aparecen?

La solución:

Protocolo	IP origen	IP destino	Puertos implicados	Descripción
IRC	10.0.2.1 5	77.243.185.106	1030 ← 6667	Conexión al servidor IRC de Undernet
FTP	10.0.2.1 5	192.168.0.52	1031, 1034 → 21	Control de sesión FTP
FTP-DATA	10.0.2.1 5	192.168.0.52	6045, 6047, 6052...	Transferencias de ficheros vía FTP pasivo

Dame un listado de las direcciones IP origen y destino que aparecen por cada protocolo

Perfecto 

Del análisis del fichero **captura.txt**, te detallo las **direcciones IP origen y destino** observadas, clasificadas **por protocolo de red** identificado en la captura:

1. IRC (puerto 6667 – protocolo de chat)

Comunicación entre el equipo local y un servidor IRC externo.

Rol	Dirección IP	Puerto	Descripción
Origen	10.0.2.15	1030	Cliente IRC local
Destino	77.243.185.106	6667	Servidor IRC (Manchester.UK.Eu.UnderNet.org)

- Comunicación bidireccional entre cliente (10.0.2.15) y servidor (77.243.185.106).

2. FTP (puerto 21 – control de sesión)

Sesión FTP entre el cliente local y un servidor en la red interna.

Rol	Dirección IP	Puerto	Descripción
Origen	10.0.2.15	1031, 1034	Cliente FTP
Destino	192.168.0.52	21	Servidor FTP (FileZilla Server 0.9.41 beta)

- Comandos observados: USER admin, PASS hack123day, CWD keylogs, STOR ...

3. FTP-DATA (canales pasivos)

Conexiones de datos iniciadas por el cliente a los puertos pasivos del servidor (modo PASV).

Rol	IP Origen	IP Destino	Puerto servidor (PASV)	Descripción
Cliente	10.0.2.15	192.168.0.52	6045	Transferencia de Keys_Apr_30_2012_16_03_50.html
Cliente	10.0.2.15	192.168.0.52	6047	Transferencia de App_Apr_30_2012_16_06_10.html
Cliente	10.0.2.15	192.168.0.52	6052	Transferencia de ano_comunicado.html (texto largo HTML)

- Todas las conexiones de datos van desde el cliente 10.0.2.15 hacia el servidor interno 192.168.0.52 en puertos dinámicos asignados por el modo PASV.

Para realizar el mismo ejercicio con el LLM privado se trata de proporcionar una URL con el contenido alojado en github público y también en pastebin.

La conclusión es que el LLM local, no está preparado para hacer el análisis por los siguientes motivos:

- Entrenamiento
- Memoria
- Rendimiento

3. ¿A qué sala se está accediendo?

Tras solicitar a que sala está accediendo la IP origen analizada del fichero captura.txt, el LLM responde lo siguiente : Se está conectando la IP origen 10.0.2.15 al canal: winsock

```
14:09:07.934343 IP 10.0.2.15.1030 > 77.243.185.106.6667: Flags [P.], seq 104:118, ack 5021, win 63688, length 14
```

```
... JOIN #Winsock
```

```
14:09:07.985647 IP 77.243.185.106.6667 > 10.0.2.15.1030: Flags [P.], seq 5021:5258, ack 118, win 65535, length 237  
:Manchester.UK.Eu.UnderNet.org 353 Anon64bits = #WinSock :Anon64bits CriMeS sold @kdcs z3n
```

The screenshot shows a dark-themed web interface for network analysis. At the top, there's a purple button labeled 'Obtener Plus'. Below it, a message asks if the user wants a network diagram. A callout box contains the text: "Quiero obtener del análisis del fichero captura.txt, a que sala IRC se estaba conectando la IP origen". The main content area displays network traffic logs in YAML format. One log entry is highlighted: "14:09:07.934343 IP 10.0.2.15.1030 > 77.243.185.106.6667: Flags [P.], seq 104:118, ack 5021 ... JOIN #Winsock". Another entry follows: "14:09:07.985647 IP 77.243.185.106.6667 > 10.0.2.15.1030: Flags [P.], seq 5021:5258, ack 11 :Manchester.UK.Eu.UnderNet.org 353 Anon64bits = #WinSock :Anon64bits CriMeS sold @kdcs z3n". At the bottom, there's a section titled "Siquieres, puedo:" with a list of options:

- extraer todas las líneas relacionadas con IRC del fichero,
- o crear un pequeño informe mostrando timestamps, comandos IRC y participantes. ¿Cuál prefieres?

4. ¿Algún servidor? En caso afirmativo, indica la versión de la aplicación.

Se le solicita al LLM, que nos proporcione la versión del servidor o aplicativo al que se accede en el tráfico del fichero captura.txt:

Dime a que servidor IRC se ha conectado informándome de la versión de la aplicación

Excelente pregunta 🌟

Del análisis del fichero **captura.txt**, se puede determinar con claridad el **servidor IRC** al que se conecta la IP origen **10.0.2.15** y la versión del software que ejecuta.

💡 **Servidor IRC identificado**

El cliente (**10.0.2.15**) se conecta al servidor:
Manchester.UK.Eu.UnderNet.org

🧠 **Detalles técnicos del servidor**

En el log del fichero se observa la siguiente respuesta del servidor después de la autenticación:

```
ruby Copiar código
`g 001 Anon64bits :Welcome to the UnderNet IRC Network, Anon64bits
`g 002 Anon64bits :Your host is Manchester.UK.Eu.UnderNet.org, running version u2.10.12.14
`g 003 Anon64bits :This server was created Thu Nov 24 2011 at 16:38:36 GMT`
```

Se le solicita un tabla específica que recopile los datos:

Campo	Valor
Servidor IRC	Manchester.UK.Eu.UnderNet.org
Red IRC	UnderNet
Versión del software	u2.10.12.14
Puerto usado	6667 (típico de IRC sin cifrado)
Fecha de creación del servidor	24 de noviembre de 2011, 16:38:36 GMT

5. ¿Hay información sensible en la captura como, por ejemplo, un login?

Se solicita información relevante al LLM, como credenciales u otros datos de interés:

Busca información sensible en el tráfico como autenticación y credenciales

Perfecto — he buscado en el contenido del fichero `captura.txt` y esto es lo sensible que se puede extraer **ahora mismo** (credenciales, rutas y otros datos que deberían considerarse comprometidos). Incluyo la evidencia extraída del propio fichero.

1) Credenciales / autenticación en claro (más importantes)

- Usuario FTP: `admin`
Evidencia: `USER admin`.
- Contraseña FTP en texto plano: `hack123day`
Evidencia: `PASS hack123day`.
- Respuesta de login FTP confirmando acceso: `230 Logged on`. (indica que las credenciales funcionaron).

2) Rutas / ficheros transferidos (posible exfiltración de datos)

- Directorio remoto accedido: `/keylogs` (CWD keylogs).
- Ficheros subidos (STOR):
 - `Keys_Apr_30_2012_16_03_50.html` (STOR).
 - `App_Apr_30_2012_16_06_10.html` (STOR).
 - Varias transferencias tipo `Transfer OK` (226 Transfer OK) que confirman finalización de transferencias.

Comentario: los nombres (`keylogs`, `Keys_*`) y el contenido HTML visible (bloques de texto que parecen logs/mensajes/comunicados) sugieren que se han transferido archivos con información sensible o registros de actividad.

3) Información de IRC que puede ser relevante (revelación de identidad/host)

- Nick utilizado en IRC: `Anon64bits` .(comandos `NICK Anon64bits`, `USER Anon64bits ...`).
- Usuario/host revelado por el servidor: `Anon64bits=+~Anon64bit@82.158.0.120.dyn.user.ono.com` .(esta línea revela un hostname / IP pública dinámica ligada al usuario).
- Canal conectado: `#WinSock` (JOIN #Winsock / NAMES list mostrando `#WinSock`).

Comentario: IRC en puerto 6667 es texto sin cifrar; cualquier nick, host, mensajes y posibles enlaces compartidos pasan en claro.

Una recopilación en formato de table extraída de los datos:

Tipo	Valor encontrado	Evidencia / contexto (extracto)
------	------------------	---------------------------------

FTP usuario	admin	USER admin en la sesión FTP hacia 192.168.0.52:21.
FTP contraseña	hack123day	PASS hack123day — se envió en texto plano y servidor respondió 230 Logged on.
FTP directorio	/keylogs	CWD keylogs y 250 CWD successful. "/keylogs" is current directory.
Ficheros subidos (STOR)	Keys_Apr_30_2012__16_03_50.html, App_Apr_30_2012__16_06_10.html, Keys_Apr_30_2012__16_10_02.html, ...	Líneas STOR <filename> y respuestas 150 Connection accepted / 226 Transfer OK.
IRC nick	Anon64bits	NICK Anon64bits y respuestas del servidor.
IRC canal	#WinSock	JOIN #WinSock y lista de nombres que muestra #WinSock.
IRC servidor (host)	Manchester.UK.Eu.UnderNet.org	Respuestas :Manchester.UK.Eu.UnderNet.org 001/002/... indicando el servidor y versión u2.10.12.14.
IPs locales / remotas detectadas	10.0.2.15, 77.243.185.106, 192.168.0.52	Aparecen en los encabezados IP <origen> > <destino>: ... de muchas tramas.
Hostmask público	Anon64bits=+-Anon64bit@82.158.0.120.dyn.user.ono.com	Línea 302 del servidor con hostmask que revela IP/host dinámico.
Indicadores "keylog" / nombres de carpeta	keylogs, Keys_*	Nombre de directorio remoto /keylogs y ficheros Keys_... — fuerte indicio de keylogger / recolección de teclas.

6. ¿Alguna carpeta dónde se pudo enviar algo?

Se solicita la carpeta y se identifica un directorio **/keylogs**:

Ficheros interesantes subidos en la carpeta, ¿es posible ver su información?

Pensó durante 13s >

Sí — he extraído del PCAP (el `captura.txt`) el contenido que se estaba enviando en las transferencias FTP y te muestro aquí lo relevante por cada fichero subido en `/keylogs`. Incluyo evidencias textuales extraídas del fichero con su referencia.

[Captura](#) [Captura](#) [Captura](#)

Ficheros detectados en `/keylogs` (subidos vía STOR)

1. `Keys_Apr_30_2012_16_03_50.html` — contenido HTML con registros de actividad y títulos de aplicaciones ejecutadas (wireshark, notepad, mIRC...). [Captura](#)

Extracto (payload capturado durante la transferencia de datos FTP):

```
php-template Copiar código

<HTML><HEAD><STYLE>...</STYLE></HEAD><META http-equiv=Content-Type content="text/html
<H1>lunes, 30 de abril de 2012 [16:00] wireshark.exe: AMD PCNET Family Ethernet Adapte
<H1>lunes, 30 de abril de 2012 [16:03] notepad.exe: Sintulo - Bloc de notas</H1>
<H1>lunes, 30 de abril de 2012 [16:05] mirc.exe: mIRC</H1>
Ya est.. el comunicado y todo, preparado
...
</BODY></HTML>
```

(Evidencia en la captura: STOR `Keys_Apr_30_2012_16_03_50.html` y segmentos del HTML).

7. ¿Puedes ver algún fichero (ya sea contenido o nombre del fichero) que pueda ser jugoso?

Se pide que analice los ficheros alojados en el directorio:

The screenshot shows a digital forensic analysis interface. At the top, there is a list of files with their paths and descriptions. Below this, a specific file is expanded to show its contents in a code editor-like window. The code editor has a dark theme, with syntax highlighting for PHP-like code. It includes a toolbar with buttons for copy, cut, and paste, and a status bar at the bottom.

2. App_Apr_30_2012_16_06_10.html — otro HTML con listado de procesos/actividades y fragmentos de texto (similar formato de informe). Fue subido inmediatamente después con STOR

App_Apr_30_2012_16_06_10.html. Captura

Extracto:

```
php-template Copiar código

<HTML>...<BODY>
<H1>lunes, 30 de abril de 2012 [16:06] : C:\WINDOWS\system32\wuauctl.exe: Actualizació
<H1>lunes, 30 de abril de 2012 [16:00] : C:\Archivos de programa\Wireshark\wireshark.e
<H1>lunes, 30 de abril de 2012 [16:00] : C:\Archivos de programa\mIRC\mirc.exe: mIRC</
...
</BODY></HTML>
```

(Evidencia: STOR App_Apr_30_2012_16_06_10.html y bloques de payload capturados). Captura

3. Keys_Apr_30_2012_16_10_02.html — hay otra transferencia posterior con nombre parecido (Keys_Apr_30_2012_16_10_02.html), también con HTML. Captura

4. Un fichero largo tipo comunicado / ano_comunicado (aparece como contenido HTML en varias tramas): contiene un manifiesto o comunicado en español que menciona detenciones relacionadas con Anonymous, ciudades, nombres parciales/seudónimos y descripciones extensas — parece ser el cuerpo principal del comunicado que también fue subido. Extractos muestran párrafos largos con referencias a detenciones, nombres, lugares y llamadas a la acción. Captura Captura

Qué información contienen (resumen y por qué son “interesantes”)

- Metadatos / cronología: los HTML contienen timestamps (30 de abril de 2012, horas alrededor de 16:00–16:11) y listados de aplicaciones abiertas (Wireshark, mIRC, Notepad, etc.). Esto sugiere que los archivos fueron generados por un sistema que iba registrando actividad de usuario/procesos.
 captura
 - Contenido potencialmente sensible:
 - Textos tipo “ano_comunicado” incluyen nombres, lugares y referencias a detenciones de personas vinculadas a Anonymous; podrían contener datos personales o incriminatorios.
 captura
 - Los ficheros llamados `Keys_*` y la carpeta `/keylogs` sugieren que se están almacenando **registros**、“**keylogs**” (posible captura de teclas o logs sensibles). Esto es crítico: nombres de ficheros y contenido HTML muestran registros y textos que fácilmente pueden contener secretos, credenciales o información exfiltrada.
 captura
 - Evidencia de transferencia exitosa: respuestas FTP `150 Connection accepted` y `226 Transfer OK` confirman que las subidas terminaron correctamente.
 captura  captura

Se le solicita que componga un fichero con la información interesante extraída, Y genera un par de ficheros HTML a los que se ha tenido acceso

Tras analizar el contenido del fichero, se genera la sospecha de que se trata de un posible keylogger, por lo que se le plantea la hipótesis al LLM :

El contenido del fichero HTML parece que proviene de un Keylogger

Sí — por lo que muestras y por lo que he extraído del HTML reconstruido, hay **evidencias claras** de que los ficheros subidos provienen de un *keylogger*, como mínimo, de un recogedor de actividad de sistema (logs de procesos, ventanas y textos). Te explico con detalle lo que he comprobado y los siguientes pasos recomendados.

Evidencia que apoya la hipótesis “keylogger / registro de actividad”

- **Nombre de carpeta remoto:** `/keylogs` — nomenclatura clara de almacenamiento de “keylogs” (registros de teclas).
- **Nombres de fichero:** `Keys_Apr_30_2012_16_03_50.html`, `Keys_Apr_30_2012_16_10_02.html` — “Keys” sugiere captura de teclas.
- **Contenido del HTML:** contiene listados de procesos y ventanas abiertas con timestamps (ej.: `wireshark.exe`, `mirc.exe`, `notepad.exe`, rutas `C:\WINDOWS\system32\...`) y grandes bloques de texto que parecen transcripciones o “comunicados” (posible texto pegado o capturado). Esto es típico de keyloggers o agentes de monitoring que registran actividad de aplicaciones y texto visible en ventanas.
- **Transferencia por FTP a servidor interno:** los archivos se suben a `192.168.0.52` (carpeta `/keyLogs`) mediante FTP en claro — indica exfiltración de los registros capturados.
- **Metadatos temporales:** timestamps en los HTML (30 de abril de 2012, horas alrededor de 16:00-16:11) que muestran el momento de la captura.
- **Otros Indicios:** la presencia de textos largos (manifiesto/comunicado) y nombres de usuarios/hosts en la captura IRC (p. ej. `Anon64bits` y hostmask `82.158.0.120...`) sugieren que se registró contenido sensible.



En conjunto, esto es bastante consistente con actividad de un keylogger/monitor que recolecta inputs y

De forma afirmativa el LLM, opina que las evidencias apoyan la hipótesis de que los ficheros HTML, del directorio /keylogs parecen ser los logs almacenados en el servidor de la ejecución de un keylogger.