


PREGUNTAS CHAT GPT HASTA ALCANZAR SOLUCIÓN FINAL

Introducción

Durante el desarrollo del proyecto de detección de fraude en Wallapop (categoría smartphones), utilizamos ChatGPT como una herramienta de apoyo para resolver dudas, entender conceptos relacionados con Elasticsearch, Kibana, Fleet y el diseño del poller, y para estructurar el código paso a paso. El objetivo de este documento es mostrar cómo interactuamos con ChatGPT, incluyendo todas las preguntas que le realizamos y las respuestas recibidas, reflejando el proceso de aprendizaje y desarrollo que nos permitió llegar al código final del poller, la ingesta de datos y los dashboards de detección de fraude.

1- Elección del tema más adecuado

 Assignment 2 NM course.pdf
PDF

tengo que hacer este trabajo, sobre que categoria lo harías? cual crees que es la categoria sobre la que se producen mas estafas? cual es la mas facil de detectar?

Yo este trabajo lo haría sobre **smartphones** ("Phones and Tablets → Smartphones").

Te explico el porqué y de paso respondo a las otras dos preguntas:

1. ¿Sobre qué categoría lo harías?

Recomendación principal: **Smartphones**

Ventajas para tu trabajo:

- **Muchísimo volumen de datos** → es fácil que con un par de días de poller tengas cientos de anuncios, lo

2- Creación fichero poller completo

vamos a empezar por el poller

Pensó durante 3m 20s >

Genial, vamos con el **poller de smartphones** 🍌

La idea (según el enunciado) es que el poller haga todo esto: llamar a la API de Wallapop con filtros (keywords, `category_id`, `time_filter=today`), procesar el JSON, añadir campos de enriquecimiento (incluido un `risk_score`) y escribir un **fichero diario** `wallapop_smartphones_YYYYMMDD.json` con un **JSON por línea**.

Te dejo directamente un `poller.py` **completo**, pensado para tu categoría **smartphones**.

3- Elastic desde máquina virtual o descargar.

pero eso se puede hacer desde un elatic de una maquina virtual? o me tengo que descargar elastic?

Pensó durante 20s >

Buena duda, porque aquí es fácil liarse 😊

1. ¿Tengo que descargar Elastic en mi portátil?

No.

Si ya estás viendo Kibana en `http://192.168.153.2:5601` dentro de la máquina virtual, significa que:

- Elasticsearch **ya está instalado** en esa VM.
- Kibana **ya está instalado** en esa VM.
- Y casi seguro que el **Elastic Agent** también (lo usasteis en el lab de IPFIX).

👉 Todo lo de Fleet + Custom Logs se hace **usando ese Elastic de la máquina virtual**, no tienes que instalar nada en tu ordenador.

4- Conectar elastic con el json

como le digo a elastic donde esta mi json

Pensó durante 29s >

Buenísima pregunta, porque aquí está el truco: **Elasticsearch no “ve” tu JSON en el disco por sí solo.**

Se lo tienes que **mandar tú** de una de estas dos formas:

- con un **script en Python (Bulk API)** → Opción A.1 del guion
- o configurando **Fleet / Custom Logs** y poniendo la **ruta del fichero** allí → Opción B

Te explico las dos, pero te dejo ya hecho el caso más sencillo (Python Bulk).

5- Guarda los json en un directorio que pueda acceder

`poller.py`
Python

este es mi poller, como lo modifico para que se me guarden los json en un sitio que elatic pueda ver

Buenísimo, ahora sí que vamos directos al grano 📌

Tu `poller.py` ahora mismo guarda el JSON en el directorio actual (donde ejecutes el script): `poller`

```
python
Copiar código

def save_items(items):
    today = datetime.utcnow().strftime("%Y%m%d")
    filename = f"wallapop_smartphones_{today}.json"

    with open(filename, "w", encoding="utf-8") as f:
        for item in items:
```

6- Comprobación json

ya he creado el json en la maquina elastic, pero ya tenia la configuracion de elastic hechi, que tengo que hacer ahora para comprobar que esta bien?

Pensó durante 13s >

Perfecto, ahora que el JSON ya está en `/var/log/wallapop` de la máquina *elastic* y la integración está creada, toca solo **comprobar que Elastic Agent se lo está tragando** y que tú lo ves en Kibana.

Te doy los pasos en orden.

1 Ver que el agente está bien y usando esa policy

1. En Kibana ve a **Fleet** → **Agents**.
2. Haz clic en el agente que aparece.
3. Comprueba:



- **Status:** debería estar `Healthy` / `Online`.

7- Corregir problemas de visualización en elastic.

 **poller2.py**
Python

he hecho este poler y lo he puesto a ejecutar con la aplicaion tmux para que me saque nuevos anuncios cada mediahora y los añada al json de ese dia. Cunado voy a elastic, en el discover solo me salen los de ayer y los de hoy no, pero en el json de hpy si que estan con todas las actualizaciones y en ese directorio. Porque?

Parece que el poller está funcionando bien para agregar nuevos anuncios al JSON del día, pero si en **Discover** no ves los anuncios de hoy, aunque estén en el archivo, es posible que haya un problema con la **lectura de los nuevos datos** por parte de Elasticsearch. Aquí te dejo algunas cosas a revisar:

1 Revisa que los nuevos anuncios están en el directorio correcto

El poller guarda los anuncios en el archivo con nomb ↓ asado en el día:

python

 Copiar código

8- Descarga de elastalert.

quiero que ahora me indiques como hacer el punto 9 paso a paso, con lo que me tengo que descargar y todo



Pensó durante 55s >

Perfecto, ahora vamos con la **Sección 9 (Elastalert2)** paso a paso, incluyendo qué instalar y qué ficheros crear.

Voy a asumir que vas a instalar Elastalert2 **en la misma máquina donde está Elasticsearch** (la VM "elastic"), y que tu índice es el data stream de Fleet tipo `logs-wallapop*`. Ajusta nombres si usas otro.

0. Qué tienes que tener ya hecho

Antes de ponerte con el punto 9, deberías tener:

- Elasticsearch y Kibana funcionando.
- Tus JSON ya se están ingiriendo en el data stream (por ejemplo `logs-wallapop-default`).
- En los documentos existen los campos:
 - `price.amount`
 - `enrichment.risk score`



9- Errores en elastalert.

I > ChatGPT 5.1 ▾

Perfecto, he analizado tu `config.yaml` y tus reglas, y aquí tienes el diagnóstico completo:

✗ POR QUÉ NO FUNCIONA TU ELASTALERT2

Hay **4 fallos críticos** que explican que Elastalert2 no lea ningún anuncio y escupa errores.

! PROBLEMA 1 — Tu índice NO es `logs-wallapop*`

Tu poller genera archiyps dentro de:

```
swift
```

Copiar código

```
/var/log/wallapop/wallapop_smartphones_YYYYMMDD.json
```

Pero **NO** veo que estés usando Elastic Agent + Fleet, por lo que no se creará un data stream llamado `logs-wallapop*`.

Resultado:



Tus reglas están apuntando a un índice que **NO** existe, por eso Elastalert2 no lee nada.

+ Pregunta lo que quieras

