**Módulo: Análisis de cabeceras X-\* (X-Headers)**

**Descripción general**

El módulo **TFM Email X-Headers** implementa una extensión para Autopsy, desarrollada en **Jython**, cuya finalidad es **extraer y analizar las cabeceras no estándar** de los mensajes de correo electrónico en formato .eml. Estas cabeceras, conocidas como *X-Headers*, son añadidas por los clientes de correo, servidores intermedios o herramientas de exportación, y constituyen una fuente de información altamente valiosa para la **adveración forense de la procedencia y autenticidad** de un mensaje.

Entre las más habituales se encuentran:

* X-Mailer: indica el cliente o aplicación que generó el mensaje (por ejemplo, *Microsoft Outlook* o *Apple Mail*).
* X-Exported-By: se utiliza cuando el correo ha sido **exportado desde un cliente** (por ejemplo, al guardar un mensaje desde Outlook a un archivo .eml).
* X-MimeOLE, X-Originating-IP y las cabeceras X-MS-\* o X-Microsoft-\*, propias de entornos Exchange o Microsoft 365.

El módulo realiza una **extracción sistemática** de todas las cabeceras que comiencen por “X-”, las decodifica correctamente (en caso de contener valores MIME codificados) y genera un **resumen analítico** con las cabeceras encontradas, su posible significado y un sistema de **puntuación de riesgo o relevancia forense**.

**Funcionamiento interno del código**

El módulo se integra en el flujo de ingestión de Autopsy mediante la clase FileIngestModule.  
Durante el análisis de cada archivo .eml:

1. **Lectura y parseo del mensaje:**  
   Se leen los bytes del archivo y se construye un objeto email.message utilizando la biblioteca estándar de Python.  
   Este proceso permite acceder a todas las cabeceras del mensaje y a su estructura MIME.
2. **Extracción de cabeceras X-**\*  
   Mediante la función core.extract\_x\_headers(msg), se recorren todas las cabeceras del mensaje y se recopilan aquellas que comienzan por “X-”.  
   Cada una se almacena como un par nombre–valor, decodificando correctamente los encabezados MIME.
3. **Análisis y heurísticas:**  
   La función core.summarize\_x\_headers\_findings() interpreta las cabeceras encontradas y calcula una **puntuación heurística** basada en su relevancia:
   * +40 puntos si existe X-Exported-By, por ser un indicio directo de que el correo **fue exportado desde un cliente** y no extraído del servidor original.
   * +10 puntos si hay X-Mailer, útil para identificar el software o dispositivo de origen.
   * +5 puntos si se detectan cabeceras pertenecientes al ecosistema Microsoft/Exchange.  
     Esta puntuación se almacena como atributo del artefacto en Autopsy.
4. **Publicación en Autopsy:**  
   Finalmente, se crea un artefacto de tipo TSK\_INTERESTING\_FILE\_HIT en el **Blackboard**, dentro del conjunto “X-Headers analysis”.  
   Este artefacto incluye una descripción textual del análisis (todas las cabeceras detectadas, valores relevantes y señales de exportación) y la puntuación calculada.

**Utilidad en el proceso de adveración**

La importancia de este módulo dentro del TFM radica en que las cabeceras X-\* son una **fuente de evidencia técnica sobre el origen, manipulación o exportación** de un correo electrónico.

En un contexto forense, los correos exportados manualmente desde un cliente pueden haber sido **modificados, re-guardados o alterados** antes de su entrega como prueba. La presencia de cabeceras como X-Exported-By o patrones característicos de clientes de correo permiten inferir que el archivo .eml **no proviene directamente del servidor**, sino de una acción manual de exportación.

Asimismo, la identificación del cliente o servidor (X-Mailer, X-MS-Exchange-Organization-\*, etc.) facilita **establecer la coherencia** entre la estructura técnica del mensaje y lo que el usuario declara sobre su procedencia. Por ejemplo, un correo supuestamente enviado desde Gmail que presenta cabeceras X-Microsoft-Exchange constituiría una **incongruencia significativa** en la cadena de custodia digital.

**Conclusión**

El módulo TFM Email X-Headers amplía la capacidad analítica de Autopsy en el ámbito del correo electrónico, introduciendo una capa de verificación centrada en **metadatos no estándar**.  
Gracias a su diseño modular y ligero, puede aplicarse automáticamente sobre grandes volúmenes de correos .eml, ofreciendo indicadores rápidos sobre **posibles manipulaciones, exportaciones o incongruencias** en los mensajes analizados.

Este componente resulta esencial para el objetivo global del TFM: **desarrollar herramientas automáticas de adveración digital de evidencias de comunicación**, garantizando tanto la integridad técnica como la trazabilidad probatoria de los mensajes electrónicos.