**Ingeniería de Software**

**PROYECTO WEB SCRAPER**

**REQUERIMIENTOS, CASOS DE USO Y DIAGRAMAS**

Mas información para añadir aquí.

**Contenido**

[**Requerimientos Funcionales** 3](#_Toc201168701)

[Requerimientos Principales 3](#_Toc201168702)

[Requerimientos Secundarios 3](#_Toc201168703)

[**Requerimientos No Funcionales** 3](#_Toc201168704)

[**Diagrama de Casos de Uso** 4](#_Toc201168705)

[**Descripción de los casos de uso** 4](#_Toc201168706)

[Realizar Scraping 4](#_Toc201168707)

[Filtro de Resultados 5](#_Toc201168708)

[Análisis de Seguridad 5](#_Toc201168709)

[**Diagramas de Actividad** 5](#_Toc201168710)

[Realizar Scraping 5](#_Toc201168711)

[Filtro de Resultados 6](#_Toc201168712)

[Análisis de Seguridad 6](#_Toc201168713)

[**Diagramas de Secuencia** 7](#_Toc201168714)

[Realizar Scraping 7](#_Toc201168715)

[Filtro de Resultados 7](#_Toc201168716)

[Análisis de Seguridad 8](#_Toc201168717)

[**Diagramas de Robustez** 8](#_Toc201168718)

[Realizar Scraping 8](#_Toc201168719)

[Filtro de Resultados 9](#_Toc201168720)

[Análisis de Seguridad 9](#_Toc201168721)

# **Requerimientos Funcionales**

## Requerimientos Principales

1. **RF-001**: El sistema debe poder extraer títulos (etiquetas <h1> a <h6>) de páginas web especificadas por el usuario.
2. **RF-002**: El sistema debe generar un índice estructurado de todos los títulos encontrados.
3. **RF-003**: El sistema debe proporcionar una interfaz interactiva para visualizar los resultados del scraping.
4. **RF-004**: El sistema debe realizar un análisis básico de seguridad (detección de formularios sin HTTPS, scripts externos, etc.).

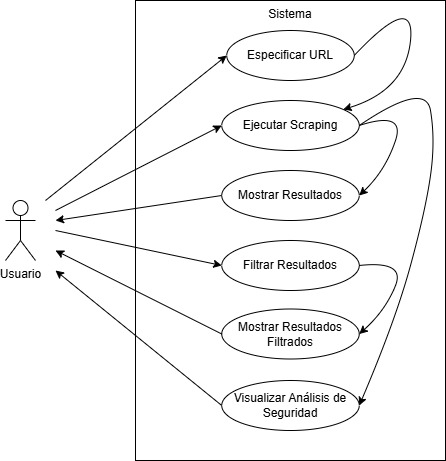
## Requerimientos Secundarios

1. **RF-006**: El sistema debe mostrar estadísticas básicas (cantidad de títulos por nivel, páginas analizadas).
2. **RF-007**: El sistema debe permitir filtrar los resultados por tipo de elemento (solo h1, solo h2, etc.).

# **Requerimientos No Funcionales**

1. **RNF-001**: El sistema debe responder en menos de 5 segundos para páginas promedio.
2. **RNF-002**: La interfaz debe ser compatible con los principales navegadores (Chrome, Firefox, Edge).
3. **RNF-003**: El sistema debe manejar páginas con hasta 10,000 elementos de título sin fallar.
4. **RNF-005**: El sistema debe respetar robots.txt y delay entre requests.

# **Diagrama de Casos de Uso**



# **Descripción de los casos de uso**

## Realizar Scraping

|  |  |
| --- | --- |
| Realizar Scraping de la Página Web | |
| Actor | Usuario |
| Descripción | El usuario ingresa una URL y el sistema extrae todos los títulos y realiza análisis de seguridad |
| Precondiciones | Conexión a internet activa |
| Flujo Principal/Alternativos | 1. Usuario ingresa URL válida 2. Sistema valida URL 3. Sistema realiza scraping 4. Sistema analiza seguridad 5. Sistema muestra resultados  * Si URL es inválida: mostrar error * Si página no accesible: mostrar error de conexión |

## Filtro de Resultados

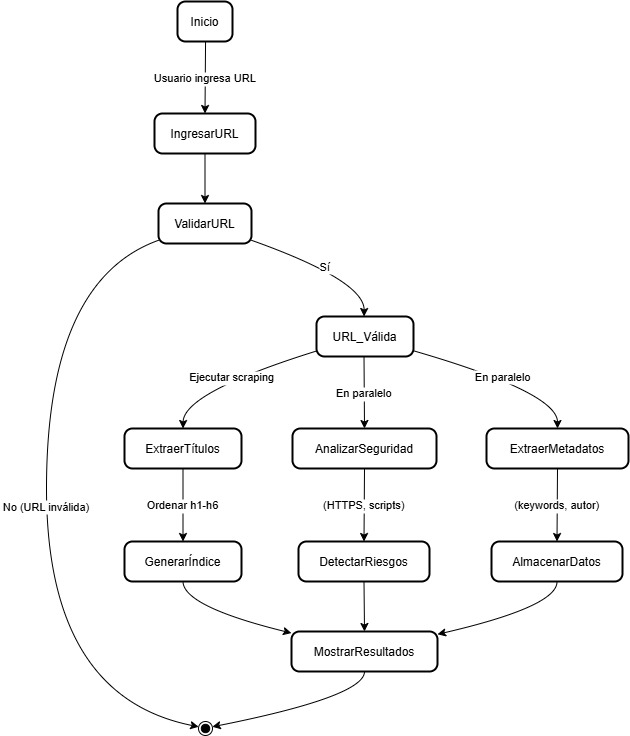
|  |  |
| --- | --- |
| Filtrar Resultados | |
| Actor | Usuario |
| Descripción | El usuario filtra los títulos mostrados por nivel (h1, h2, etc.) |
| Precondiciones | Haber realizado al menos un scraping |
| Flujo Principal/Alternativos | 1. Usuario selecciona opción de filtrado 2. Sistema aplica filtros 3. Sistema muestra resultados filtrados |

## Análisis de Seguridad

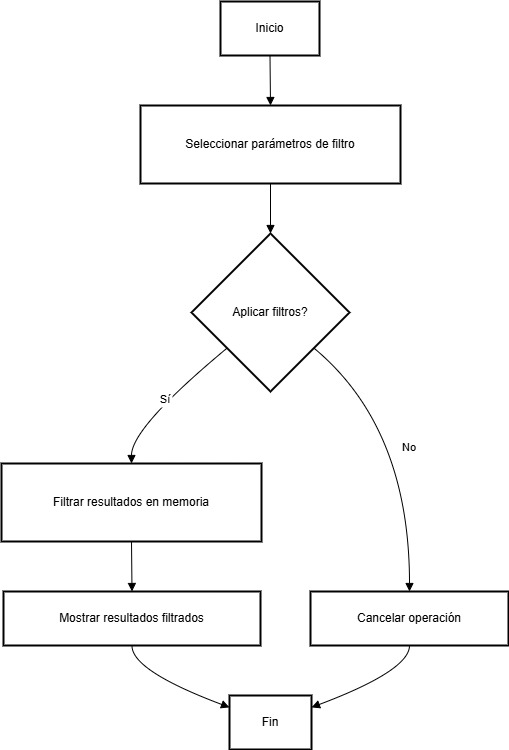
|  |  |
| --- | --- |
| Visualizar Análisis de Seguridad | |
| Actor | Usuario |
| Descripción | El usuario revisa los hallazgos de seguridad detectados |
| Precondiciones | Haber realizado al menos un scraping |
| Flujo Principal/Alternativos | 1. Usuario selecciona pestaña de seguridad 2. Sistema muestra hallazgos clasificados 3. Usuario puede expandir cada hallazgo para más detalles |

# **Diagramas de Actividad**

## Realizar Scraping



## Filtro de Resultados

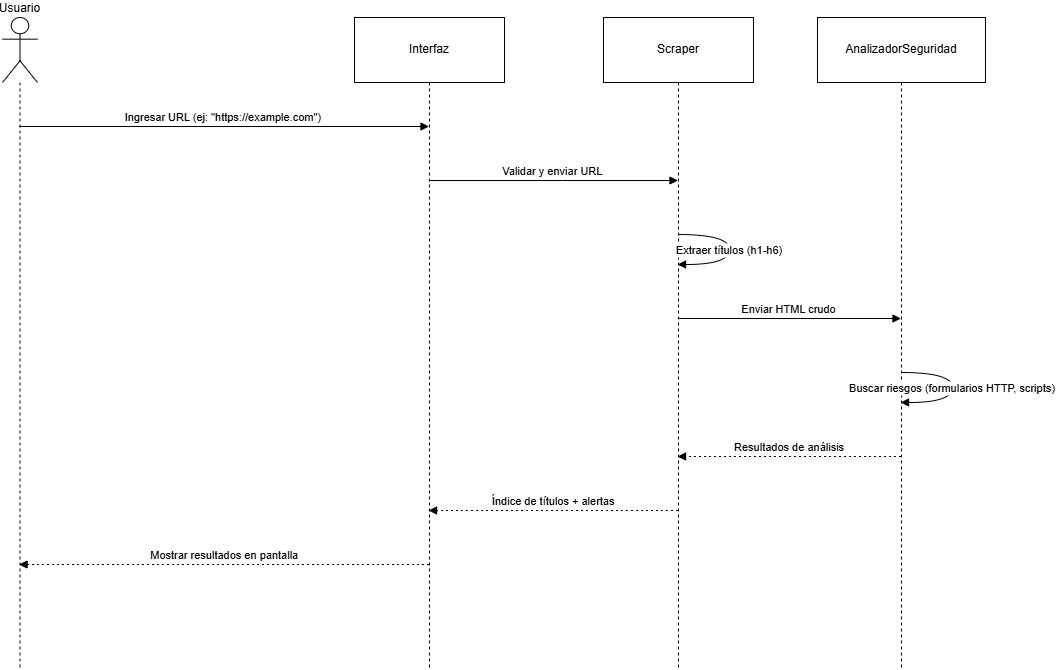


## Análisis de Seguridad

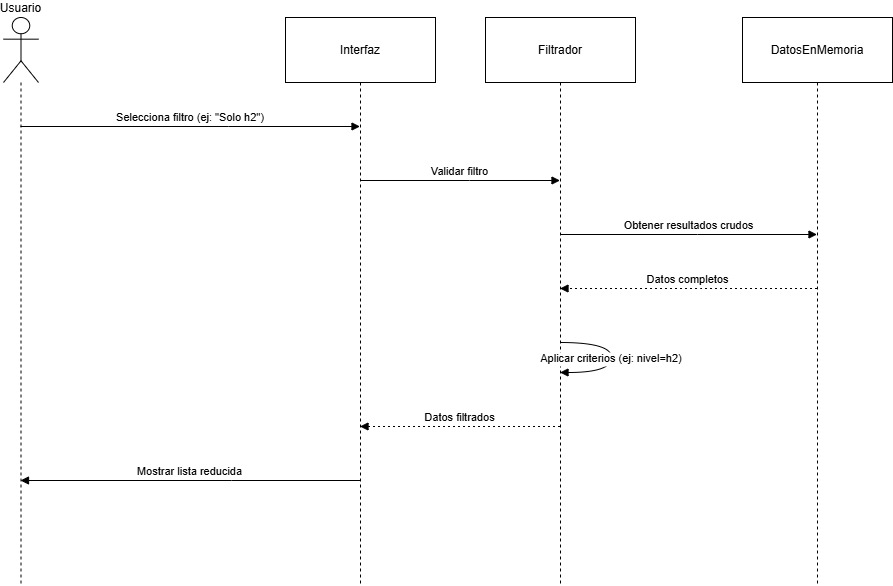


# **Diagramas de Secuencia**

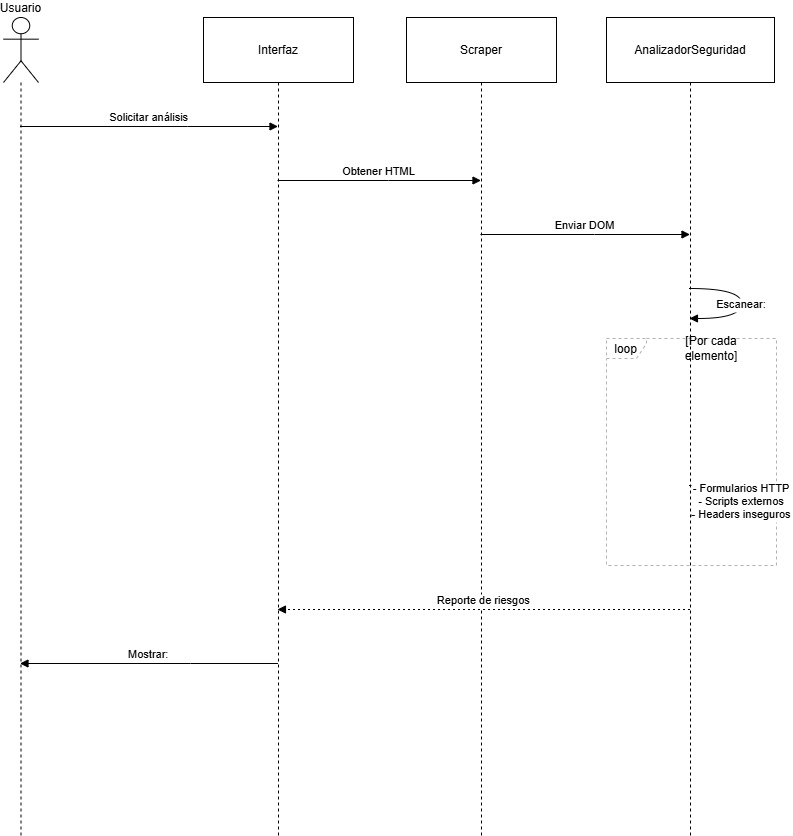
## Realizar Scraping



## Filtro de Resultados

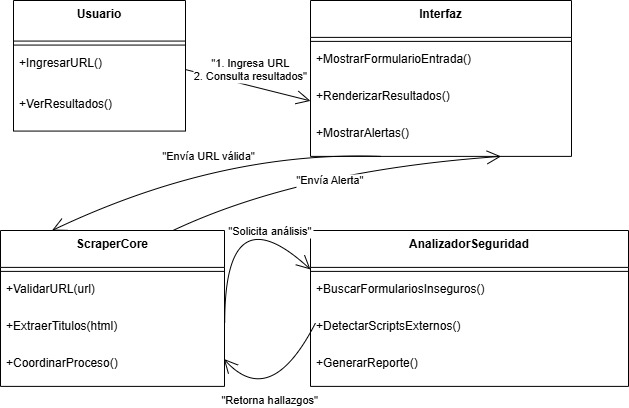


## Análisis de Seguridad

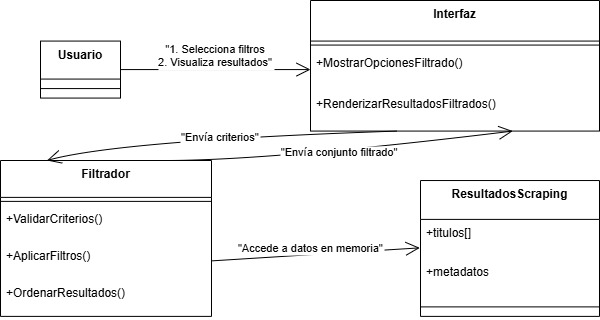


# **Diagramas de Robustez**

## Realizar Scraping



## Filtro de Resultados



## Análisis de Seguridad

