

## Ejercicio para aplicantes

### Introducción:

En Qanlex transformamos el acceso a la justicia a través del financiamiento de litigios. Nuestro propósito es identificar demandas de alto monto, contactar a los demandantes y ofrecer soluciones de financiamiento estratégico mediante la compra parcial o total de sus demandas. Para lograrlo, recolectamos información directamente de páginas web judiciales y la integramos a nuestra Plataforma Qantech, la cual organiza y filtra los expedientes según una variedad de criterios, optimizando la identificación de oportunidades clave.

### Detalle del Ejercicio:

Tu tarea es desarrollar un scraper web para la siguiente página: <http://scw.pjn.gov.ar/scw/home.seam>. El objetivo es extraer información relevante sobre demandas de presentes en el sitio.

Desarrollo del Scraper: Crear un scraper funcional que extraiga al menos: Expediente, Dependencia, Demandante, demandado, caratula, tipo de demanda, juzgado o tribunal y fechas relevantes. El scraper debe buscar "Parte", en la Jurisdicción "COM", utilizando la palabra clave "RESIDUOS". Los resultados deben ser guardados en una tabla.

Calidad del Código: Valoramos un código limpio, bien estructurado y siguiendo buenas prácticas de programación.

Documentación: Incluye comentarios y/o documentación que faciliten la comprensión y mantenimiento del código, así como cualquier cosa que hayas asumido del ejercicio.

### Puntos Adicionales (Opcionales):

Se considerará positivamente incluir cualquier de los siguientes ítems y/o sub ítems.

- 1) Captcha: La página cuenta con un captcha. Se considerará un plus si logras automatizar su resolución. Si esto no es posible, puedes implementar una pausa (time.sleep) y resolverlo manualmente durante la ejecución.
- 2) Automatización y Despliegue en AWS:
  - Crear una instancia en AWS para ejecutar el scraper.
  - Configurar la instancia para que se inicie y detenga en horarios programados.
  - Hacer que el scraper se ejecute automáticamente al iniciarse la instancia.
  - Utilizar instancias dentro de subnets privadas en AWS.
- 3) Almacenamiento de Datos:
  - Guardar los datos extraídos en una tabla MySQL para facilitar su consulta y análisis.
- 4) Control de Versiones y Despliegue Continuo:
  - Subir el código a un repositorio de GitHub.
  - Configurar la instancia de AWS para que obtenga el código actualizado desde GitHub cada vez que se inicie, o implementar un sistema de despliegue continuo que actualice el código en la instancia al realizar cambios en el repositorio.

Entrega: Envía el código fuente y cualquier documentación adicional por correo electrónico a [yago@qanlex.com](mailto:yago@qanlex.com) y [jcolombo@qanlex.com](mailto:jcolombo@qanlex.com). Si utilizas GitHub, comparte el enlace al repositorio donde se encuentra el código. Incluye instrucciones detalladas sobre cómo ejecutar el scraper y configurar el entorno si es necesario. La fecha límite para la entrega será el 21 de diciembre.