Definición Proyecto APT – Fase 1

Nombre: Paulo Irribarra

Carrera: Ingeniería en Informática

Proyecto APT: Web Scraper para comparación de precios de medicamentos

# Abstract

## English

This project proposes the development of a Web Scraper aimed at collecting and analyzing medicine prices from different pharmacies. The main purpose is to build an automated system that extracts, cleans, and stores data for comparison, providing users with a tool that improves transparency in the pharmaceutical market. The project is relevant for the software quality and automation field, since it applies good practices in data integration, automated testing, and software development, which are key skills for my professional career.

## Español

Este proyecto propone el desarrollo de un Web Scraper orientado a recolectar y analizar precios de medicamentos desde distintas farmacias. El objetivo principal es construir un sistema automatizado que extraiga, limpie y almacene datos para realizar comparaciones, entregando a los usuarios una herramienta que mejore la transparencia en el mercado farmacéutico. El proyecto es relevante para el ámbito de calidad de software y automatización, ya que aplica buenas prácticas en integración de datos, pruebas automatizadas y desarrollo de software, competencias claves para mi carrera profesional.

# Descripción del Proyecto APT

El proyecto consiste en diseñar e implementar un Web Scraper capaz de obtener información de precios y disponibilidad de medicamentos en diferentes farmacias online. Los datos recopilados serán almacenados en una base de datos estructurada, para luego ser consumidos por un módulo de análisis que permita comparaciones y generación de reportes. Su relevancia en el campo laboral radica en la creciente necesidad de automatizar procesos de extracción de información y transformar datos en conocimiento útil, apoyando la toma de decisiones tanto para usuarios como para empresas.

# Relación con las Competencias del Perfil de Egreso

- Realizar pruebas de certificación: el proyecto aplicará pruebas automatizadas para validar el correcto funcionamiento del scraper y la integridad de los datos obtenidos.  
- Gestionar proyectos informáticos: se planificará cada fase del desarrollo, estableciendo tiempos, recursos y entregables.  
- Construir modelos de datos: se diseñará una base de datos para soportar el almacenamiento y consulta de la información extraída.  
- Desarrollar soluciones de software: se construirá un sistema funcional aplicando buenas prácticas de programación, arquitectura y mantenimiento.

# Relación con Intereses Profesionales

Mis intereses profesionales se centran en el QA automatizado, el desarrollo de software y el análisis de datos. Este proyecto refleja de manera directa estos intereses, ya que combina la automatización de procesos, la validación de datos y el desarrollo de una solución realista que se puede aplicar a un caso de uso concreto.

# Factibilidad del Proyecto

El proyecto es factible de realizar en el marco de la asignatura, considerando:  
- Tiempo: se ajusta al periodo académico, con fases claras de desarrollo.  
- Recursos: solo requiere un entorno de desarrollo (Python, Java o Node.js), librerías de scraping (BeautifulSoup, Selenium o Puppeteer) y una base de datos relacional.  
- Posibles dificultades: bloqueos de sitios web o cambios en el HTML; se mitigará implementando estrategias de rotación de user agents y manejo de errores.

# Objetivos

Objetivo General: Desarrollar un Web Scraper que permita extraer y analizar precios de medicamentos en distintas farmacias online, generando reportes comparativos de utilidad para los usuarios.

Objetivos Específicos:  
1. Diseñar e implementar un sistema de scraping para la extracción de datos de sitios web seleccionados.  
2. Diseñar un modelo de base de datos que permita almacenar y consultar la información obtenida.  
3. Implementar pruebas automatizadas para validar la consistencia y exactitud de los datos extraídos.  
4. Generar reportes de comparación de precios como evidencia de los resultados del sistema.

# Metodología de Trabajo

El proyecto se desarrollará utilizando una metodología ágil (Scrum), con iteraciones cortas que permitan evaluar avances de manera continua:  
- Levantamiento de requerimientos y análisis del dominio.  
- Diseño del scraper y definición de sitios a analizar.  
- Desarrollo incremental del scraping y almacenamiento de datos.  
- Pruebas automatizadas y control de calidad.  
- Generación de reportes y evidencias finales.

# Plan de Trabajo

- Semana 1-2: Análisis de requerimientos y diseño del modelo de datos.  
- Semana 3-4: Implementación inicial del Web Scraper para farmacias seleccionadas.  
- Semana 5-6: Integración con la base de datos y almacenamiento de datos recolectados.  
- Semana 7: Desarrollo de reportes comparativos y dashboards básicos.  
- Semana 8: Pruebas, validaciones y ajustes.  
- Semana 9: Presentación final y entrega de evidencias.  
  
Facilitadores: conocimientos previos en QA, scraping y bases de datos.  
Obstaculizadores: cambios en las páginas web o bloqueos de scraping.

# Evidencias

- Código fuente del Web Scraper.  
- Scripts de base de datos y modelo diseñado.  
- Pruebas automatizadas y reportes de ejecución.  
- Reportes comparativos generados a partir de los datos.  
- Documentación técnica del proyecto.

# Conclusiones (English)

In conclusion, this project will strengthen my professional skills by combining software development, data integration, and quality assurance. It provides a real-world scenario where automation and data analysis are essential, aligning perfectly with my career goals. The Web Scraper project is not only technically feasible but also highly valuable, as it integrates several competencies from my academic training into a practical and relevant solution.

# Reflection (English)

Reflecting on my professional path, this project represents a key step in consolidating my role as a QA Automation Engineer. It allows me to apply testing strategies, software development skills, and data management techniques in a meaningful context. In the future, I aim to expand this knowledge to more complex architectures and incorporate security aspects. This project prepares me for real industry challenges and reinforces my interest in pursuing certifications such as ISTQB to continue growing in the software quality field.