PROYECTO APT - COMPARADOR DE PRECIOS CON WEB SCRAPING.

Estudiante: Marc Matheus

Carrera: Ingeniería en Informática

Asignatura: APT122 - Proyecto de Título

Institución: Duoc UC

Profesor Guía: Cristobal Camilo Beltran

Fecha: [4-09-2025]

INDICE.

[**Abstract 3**](#_heading=h.1kmkqzg93mn2)

[**1. Descripción de Proyecto APT 4**](#_heading=)

[**2. Relación del proyecto APT con las competencias del perfil de egreso 5**](#_heading=)

[**3. Relación del proyecto con tus intereses profesionales 5**](#_heading=)

[**4. Argumento de factibilidad dentro de la asignatura 6**](#_heading=)

[**5. Objetivos claros y coherentes 6**](#_heading=)

[**6. Propuesta metodológica de trabajo 7**](#_heading=)

[**7. Plan de trabajo 8**](#_heading=)

[**8. Evidencias esperadas 10**](#_heading=)

[**Reflexiones individuales (English) 11**](#_heading=h.u3mzwmirujsk)

[**Conclusiones individuales (English) 11**](#_heading=h.xp0v0wfo8ct2)

# Abstract

(Español)

Este proyecto consiste en el desarrollo de una plataforma web que compara precios de instrumentos musicales en múltiples tiendas en línea mediante técnicas de web scraping. La solución normaliza y consolida la información de precios, ofrece búsqueda por producto, historial de precios y alertas configurables. En paralelo, integra un módulo de comunidad donde los usuarios crean hilos, comparten reseñas y votan respuestas, fortaleciendo el intercambio de conocimiento entre músicos. El objetivo es reducir asimetrías de información y mejorar la toma de decisiones de compra dentro del nicho musical, a la vez que se fomenta la convivencia digital del ecosistema. El proyecto aplica competencias de desarrollo de software, bases de datos, seguridad y gestión ágil, y se implementa con tecnologías conocidas por el equipo (Python/FastAPI, React, PostgreSQL), lo que asegura su factibilidad en el periodo académico.

(English)

This project develops a web platform that compares musical instrument prices across multiple online stores using web scraping. The solution normalizes and consolidates price data, provides product search, price history, and custom alerts. In parallel, it integrates a community module where users create threads, share reviews, and upvote answers, fostering knowledge exchange among musicians. The goal is to reduce information asymmetry and improve purchase decisions while cultivating a collaborative online space for the music niche. The project leverages competencies in software development, databases, security, and agile management, and it is implemented with technologies familiar to the team (Python/FastAPI, React, PostgreSQL), ensuring feasibility within the academic term.

# 1. Descripción de Proyecto APT

El proyecto corresponde a una plataforma web especializada en instrumentos musicales, que combina un motor de búsqueda basado en web scraping con un módulo de comunidad digital.

El primer módulo, el comparador de precios, tiene como finalidad reducir la dispersión de información en el mercado de instrumentos musicales. Actualmente, los consumidores deben navegar en múltiples páginas de tiendas, cada una con interfaces distintas y sin certeza de obtener la mejor oferta. El sistema que proponemos automatiza este proceso mediante scrapers que recopilan información de varias tiendas, la normaliza en una base de datos estructurada y la presenta al usuario en un formato consolidado. El usuario podrá:

* Buscar un instrumento por nombre, categoría o marca.
* Visualizar un ranking de precios de distintas tiendas.
* Revisar un historial de precios para observar variaciones en el tiempo.
* Configurar alertas cuando un producto baje de cierto umbral.

El segundo módulo, la comunidad, busca dar respuesta a otra necesidad: el espacio de interacción entre músicos. Aquí los usuarios podrán abrir hilos de discusión, compartir reseñas sobre tiendas o productos, votar las contribuciones de otros y aprender en conjunto. Este módulo no solo fortalece la convivencia digital, sino que también agrega confianza a los datos del comparador, al complementar la información objetiva con experiencias subjetivas.

La propuesta genera un ecosistema mixto (datos + comunidad), que entrega valor agregado a PYMEs, tiendas y consumidores. Para los estudiantes del equipo, implica un reto completo en términos técnicos (scraping, backend, frontend, seguridad, despliegue) y metodológicos (Scrum, documentación, control de avances).

# 2. Relación del proyecto APT con las competencias del perfil de egreso

Desarrollo de soluciones de software: Frontend en React y backend en Fast API aplicando buenas prácticas (modularidad, pruebas, control de versiones).

Modelado y administración de datos: Diseño e implementación de PostgreSQL para productos, ofertas, historial y contenidos de comunidad; migraciones y consultas optimizadas.

Seguridad informática: Autenticación con JWT, cifrado de contraseñas, validación de entradas, protección básica ante inyección/XSS.

Gestión de proyectos informáticos: Planificación y control con Scrum (Jira), seguimiento de issues y CI/CD.

Transformación de datos: Normalización de precios y productos obtenidos por scraping; detección simple de valores atípicos.

# 3. Relación del proyecto con tus intereses profesionales

Mis intereses profesionales están enfocados principalmente en tres áreas: ciberseguridad, entendimiento de los datos y gestión ágil de proyectos. Este proyecto se conecta directamente con ellos de la siguiente manera:

Ciberseguridad: la implementación de autenticación mediante JWT, cifrado de contraseñas y validaciones contra ataques comunes (inyecciones, XSS) me permitirá aplicar medidas de protección esenciales en entornos web. El hecho de construir una plataforma abierta al público refuerza la importancia de asegurar la integridad y confidencialidad de los datos, competencias críticas en el campo profesional de la seguridad informática.

Entendimiento de los datos: el desarrollo de scrapers y la normalización de información proveniente de distintas fuentes me permitirá trabajar con datos heterogéneos y ruidosos, extrayendo valor mediante procesos de limpieza, organización y análisis. Asimismo, el historial de precios y las alertas constituyen un ejercicio práctico en la transformación de datos en información útil para la toma de decisiones.

Gestión ágil: la aplicación de la metodología Scrum en la planificación y seguimiento del proyecto fortalece mis habilidades en gestión ágil, ya que exige coordinación, priorización de tareas y comunicación efectiva entre los integrantes. Estas prácticas son fundamentales en el ámbito laboral, donde los ciclos cortos de entrega y la adaptabilidad son cada vez más valorados.

En conjunto, el proyecto me permitirá integrar estas tres áreas de interés profesional, fortaleciendo mi perfil como ingeniero en informática con una orientación hacia la seguridad, el manejo inteligente de datos y la gestión eficiente de proyectos.

# 4. Argumento de factibilidad dentro de la asignatura

Duración: el calendario permite un MVP funcional (comparador + comunidad) dentro del periodo.

Horas y equipo: tres integrantes con reparto claro (frontend/comunidad; backend/BD/seguridad; scraping/normalización).

Materiales/Tecnologías: Python (Requests/BeautifulSoup), FastAPI, React+Tailwind, PostgreSQL, Docker, Render/Vercel; todas conocidas por el equipo.

Facilitadores: alcance acotado al nicho musical y a 5–8 tiendas iniciales; arquitectura modular.

Dificultadores y mitigación: cambios en HTML de tiendas → pruebas y adapters por tienda; límites de tiempo → priorización estricta (comparador primero, comunidad como segunda prioridad); fallback con datasets de prueba si una tienda falla.

# 5. Objetivos claros y coherentes

Objetivo general

Desarrollar una plataforma web que compare precios de instrumentos musicales mediante web scraping e integre un módulo de comunidad para apoyar la decisión de compra y el intercambio de conocimiento.

Objetivos específicos:

* Implementar scrapers para ≥5 tiendas con normalización de productos y precios.
* Diseñar e implementar una BD relacional con historial de precios y contenidos de la comunidad.
* Desarrollar API REST para búsqueda, comparativas y operaciones de comunidad.
* Construir un frontend responsivo con buscador, ficha de producto e interfaz de foro.
* Incorporar autenticación (JWT) y validaciones de seguridad esenciales.
* Entregar documentación técnica, manual de usuario y demo del MVP.

# 6. Propuesta metodológica de trabajo

El proyecto se gestionará con la metodología ágil Scrum, aplicada en ciclos iterativos de 2 semanas. Se usarán herramientas de apoyo como Jira (gestión de backlog), GitHub (repositorios y control de versiones) y CI/CD (despliegue en entornos de prueba y demo).

Además de los artefactos ágiles, se incorporarán documentos de ingeniería de software que respalden el proceso y las decisiones:

* Carta Gantt: cronograma detallado de las 18 semanas académicas.
* Casos de uso: identificación de actores (usuario invitado, usuario registrado, administrador) y sus interacciones principales.
* Diagrama de arquitectura: descripción de los módulos (scraping, API, frontend, comunidad).
* Modelo entidad-relación: estructura de la base de datos en PostgreSQL.
* Documento de definición del proyecto (fase 1).
* Memoria técnica (fase final).
* Manual de usuario (fase final).

De esta forma, la metodología no solo asegura entregables funcionales, sino también trazabilidad documental, requisito clave en el ámbito académico y profesional.

Roles:

Marc (Frontend, Comunidad & Seguridad): Responsable del frontend (UI/UX) en React y Tailwind, desarrollo del módulo de comunidad (foros, reseñas, votaciones) y además de la implementación de medidas de ciberseguridad a nivel de frontend y backend (validaciones, pruebas de seguridad, control de roles, encriptación de datos). También se encargará del manual de usuario y documentación relativa al uso seguro del sistema.

Javier (Backend & Base de datos): Encargado del backend con FastAPI, construcción de APIs REST, diseño y mantenimiento de la base de datos PostgreSQL, así como la integración con el scraping y el despliegue en entornos de prueba/producción.

Christian (Scraping & Datos): Responsable del desarrollo de scrapers para recolectar información de precios desde múltiples tiendas, la normalización de datos obtenidos y la realización de pruebas de extracción e integración con el backend.

# 7. Plan de trabajo

| Semana | Cronograma oficial APT | Desarrollo técnico y documental | Entregables / Evidencias |
| --- | --- | --- | --- |
| 1 (11–16 ago) | Fase 1: Autoevaluación de competencias y definición inicial | Discusión de ideas; autoevaluación individual; exploración de tecnologías | Documento de autoevaluación (individual) |
| 2 (18–23 ago) | Fase 1: Desarrollo de idea, cronograma, definición de tecnologías | Selección de proyecto final (comparador + comunidad); definición de cronograma | Documento de definición preliminar |
| 3 (25–30 ago) | Fase 1: Definición de proyecto | Ajuste de alcance; preparación de presentación inicial; elección de stack tecnológico | Diagrama de arquitectura inicial; borrador casos de uso |
| 4 (1–6 sep) | Fase 1: Exposición grupal de proyecto | Presentación de la propuesta final | Presentación inicial |
| 5 (8–13 sep) | Fase 2: Desarrollo / Fortalezas y debilidades | Configuración de repositorios y entornos; modelo entidad-relación; carta Gantt formal | Repositorio GitHub, Carta Gantt |
| 6 (15–20 sep) | Fase 2: Desarrollo con conocimientos previos | Scrapers iniciales (2 tiendas); migraciones BD; API health check | Código scraping v1; BD inicial |
| 7 (22–27 sep) | Fase 2: Desarrollol | API productos y precios; frontend home y buscador | API base; UI inicial |
| 8 (29 sep–4 oct) | Fase 2: Desarrollo / Fortalezas y debilidades | Scraping completo (≥5 tiendas); ficha de producto con tabla comparativa | Comparador MVP v1 |
| 9 (6–11 oct) | Fase 2: Desarrollo | Historial de precios; alertas básicas; mejoras UI | Comparador MVP v2 |
| 10 (13–18 oct) | Fase 2: Entrega de avance | Comunidad: crear hilos y responder | Entrega parcial (MVP comparador + comunidad base) |
| 11 (20–25 oct) | Fase 2: Desarrollo / Fortalezas y debilidades | Comunidad: votos y reseñas; mejoras UI | Comunidad MVP v1 |
| 12 (27 oct–1 nov) | Fase 2: Desarrollo | Autenticación JWT; encriptación de contraseñas | Módulo seguridad |
| 13 (3–8 nov) | Fase 2: Desarrollo | Integración completa scraping + comunidad + auth | Plataforma integrada v1 |
| 14 (10–15 nov) | Fase 2: Desarrollo | Pruebas funcionales; optimización; logging | Informe de pruebas |
| 15 (17–22 nov) | Fase 2: Entrega final del proyecto | Documentación técnica; manual de usuario; demo desplegada | Entrega final |
| 16 (24–29 nov) | Fase 3: Presentación | Preparación demo y presentación | Borrador presentación final |
| 17 (1–6 dic) | Fase 3: Presentación | Exposición oficial del proyecto | Presentación final |
| 18 (examen) | Fase 3: Examen | Defensa y retroalimentación | Defensa oral |

# 8. Evidencias esperadas

Avance – Documento de definición: corresponde al texto inicial donde se presentan los fundamentos del proyecto, los objetivos generales y específicos, así como la metodología seleccionada. Su valor radica en que da cuenta del encuadre inicial del trabajo y orienta todo el desarrollo posterior.

Avance – Autoevaluación individual: cada integrante elaborará una reflexión escrita sobre sus competencias, fortalezas y debilidades en relación al perfil de egreso. Esta evidencia es importante porque permite reconocer las áreas de mayor dominio y aquellas que deben fortalecerse durante el proyecto.

Avance – Carta Gantt: se diseñará una planificación temporal detallada de las dieciocho semanas del semestre, con las tareas distribuidas y los responsables asignados. Esta evidencia facilita la visualización del cronograma y asegura un control ordenado del trabajo.

Avance – Casos de uso: se elaborarán diagramas y descripciones de los actores principales (usuario invitado, usuario registrado y administrador) y sus interacciones con el sistema. Esta evidencia garantiza claridad en los requerimientos funcionales y la trazabilidad de las funcionalidades.

Avance – Repositorio en GitHub: se mantendrá un repositorio centralizado con control de versiones, donde se registrarán los avances en el código de scraping, backend y frontend. Es una evidencia del trabajo colaborativo, la aplicación de buenas prácticas y el progreso técnico del equipo.

Final – Plataforma web funcional (MVP): se entregará un sistema desplegado que integre el comparador de precios con web scraping y el módulo de comunidad, constituyendo la evidencia central del cumplimiento de objetivos.

Final – Memoria técnica: se redactará un documento formal que contendrá la arquitectura del sistema, decisiones técnicas, diagramas de apoyo y pruebas realizadas. Esta evidencia muestra la capacidad analítica y el rigor técnico alcanzado.

Final – Manual de usuario: se preparará un texto práctico y accesible que explique cómo utilizar la plataforma, desde la búsqueda de productos hasta la interacción en la comunidad. Su propósito es facilitar la adopción del prototipo por parte de usuarios no técnicos.

Final – Presentación con demo en vivo: se realizará una exposición final ante los docentes con apoyo de diapositivas y una demostración funcional del sistema. Esta evidencia permite evaluar la comunicación, la síntesis y la capacidad de defender el trabajo realizado.

# Reflexiones individuales (English)

Developing this project has been a valuable learning experience that combined my main professional interests: cybersecurity, data understanding, and agile management. On the technical side, building a web platform that collects and processes data through web scraping challenged me to deal with unreliable and heterogeneous sources. This process improved my ability to transform raw data into meaningful information, such as price comparisons and alerts. At the same time, implementing authentication, encryption, and secure coding practices strengthened my awareness of the importance of protecting users and systems against potential threats, reinforcing my interest in cybersecurity as a professional path.

From a methodological perspective, working with Scrum allowed me to experience agile management in practice. Planning sprints, defining tasks, and delivering incremental results gave structure to our teamwork and demonstrated the value of adaptability and collaboration. Integrating all these elements in a single project showed me that innovation does not depend only on technical knowledge, but also on the ability to coordinate efforts and maintain a clear focus on user needs. This reflection motivates me to continue growing as an engineer oriented towards secure, data-driven, and agile solutions.

# Conclusiones individuales (English)

Throughout this project I was able to strengthen my technical and professional skills by focusing on the areas that interest me most: cybersecurity, data understanding, and agile management. Working on the platform gave me the opportunity to apply secure coding practices, authentication mechanisms, and data protection strategies, which deepened my perspective on the importance of building safe and reliable systems. At the same time, developing the data extraction and normalization processes helped me understand how to transform raw information into valuable insights for users. Finally, applying Scrum in our teamwork allowed me to experience agile management in practice, improving my ability to plan, prioritize, and adapt to changes. Overall, this project contributed significantly to my professional growth and reaffirmed my commitment to continue developing as an engineer focused on secure, data-driven, and agile solutions.