**PRACTICA XML**

## Introducción

La Práctica 3 de PIAT, centrada en el procesamiento de documentos XML, representa un hito significativo en el marco de este proyecto de fin de grado. En esta etapa, se aborda la implementación de un analizador de documentos XML utilizando la tecnología SAX (Simple API for XML). El objetivo principal es familiarizar a los estudiantes con SAX y fortalecer sus habilidades en el diseño de algoritmos eficientes para la extracción y transformación de información a partir de fuentes de contenidos estructurados.

## Desarrollo

La implementación de la Práctica 3 se basa en la creación de un analizador de documentos XML denominado **XMLParser**, desarrollado utilizando JavaCC, una herramienta que permite generar analizadores léxicos y sintácticos en el lenguaje Java. Este analizador se ha diseñado para procesar documentos XML que contienen datos publicados en el Portal de Datos Abiertos del Ayuntamiento de Madrid ([http://datos.madrid.es](http://datos.madrid.es/)).

Los documentos XML que se procesan a través de esta herramienta presentan ciertas características clave, como la identificación de recursos mediante URI, la asociación a categorías, la agrupación en conjuntos de datos y una clasificación jerárquica basada en características temáticas. La tarea principal del analizador es extraer información específica relacionada con códigos de categorías y, posteriormente, generar un nuevo documento XML que contiene los resultados deseados.

## Conclusiones

La Práctica 3 ha brindado a los estudiantes una experiencia valiosa en el manejo de tecnologías de análisis de documentos XML. A través del uso de JavaCC y la tecnología SAX, se ha proporcionado una base sólida para procesar información estructurada de manera eficiente. Los estudiantes han tenido la oportunidad de aplicar conceptos relacionados con la validación de argumentos, el análisis de elementos XML y la generación de documentos de resultados.

Uno de los aspectos más destacados de esta práctica es la capacidad de filtrar y seleccionar información relevante basada en códigos de categorías. Esta habilidad es fundamental en situaciones donde la extracción selectiva de datos es esencial.

## Herramienta

La principal herramienta desarrollada en esta práctica es el analizador **XMLParser**. Este analizador es capaz de llevar a cabo varias tareas esenciales:

* Validación de argumentos de entrada para garantizar que los parámetros cumplan con las especificaciones requeridas.
* Extracción de información relevante de los documentos XML, siguiendo las reglas y estructuras definidas en la gramática.
* Generación de documentos de resultados que cumplen con un esquema específico.

El analizador **XMLParser** representa una contribución significativa en términos de procesamiento y manipulación de documentos XML.

## Preguntas Frecuentes

**Pregunta 1:** ¿Cuál es la tecnología utilizada para procesar los documentos XML en esta práctica?

**Respuesta:** La tecnología utilizada es SAX (Simple API for XML), que proporciona un enfoque basado en eventos para procesar documentos XML.

**Pregunta 2:** ¿Cuál es el propósito principal de esta práctica en términos de procesamiento de información?

**Respuesta:** El objetivo principal de la Práctica 3 es la extracción selectiva de información relacionada con códigos de categorías en documentos XML, así como la generación de un nuevo documento XML que contenga los resultados deseados.

**Pregunta 3:** ¿Cómo se generó el analizador **XMLParser**?

**Respuesta:** El analizador **XMLParser** se generó utilizando JavaCC, una herramienta que permite crear analizadores léxicos y sintácticos en el lenguaje Java. Este analizador sigue reglas específicas definidas en una gramática.