



Trabajo Práctico Especial

72.32 Diseño y Procesamiento de Documentos XML

Primer Cuatrimestre de 2024

Grupo 13

Integrantes:

Jonathan Lucas Blankleder - 64660

Ignacio Gastón Frund - 61465

Agustín Germán Ramírez Donoso - 64716

Introducción

Este trabajo consistió en la recuperación de información y estadísticas de los pilotos de la competición automovilística estadounidense NASCAR de cierto año y categoría, en formato XML desde una base de datos deportiva externa mediante una API ofrecida por la misma, para luego filtrar dicha información mediante consultas XQuery y generar un documento XSL FO que finalmente convertiremos en PDF. Junto al funcionamiento interno se debieron controlar los posibles errores y se creó un script bash que genera el archivo PDF automáticamente a partir del año y categoría solicitados.

Desarrollo del trabajo

Trabajo Realizado

Adjunto a este informe se encuentra el script bash *tpe.sh* que realiza todo el proceso de consulta y presentación automáticamente a partir de un *year* y *type* apropiados pasados por parámetro. En caso de que no sean apropiados, se arrojará un error.

El mismo obtiene los archivos *drivers_list.xml* y *drivers_standings.xml* con la herramienta cURL a partir de dichos parámetros y la clave de la API de SportRadar creada específicamente para el trabajo. Además, a ambos archivos se les quita el Namespace para utilizarlos sin problemas.

Una vez que cuenta con ambos archivos, realiza la consulta *extract_nascar_data.xq*, que crea el archivo *nascar_data.xml* con la información deseada en el orden deseado para la presentación del documento.

En base a este documento XML, el archivo XSLT *generate_fo.xsl* crea el archivo *nascar_page.fo* con el estilo solicitado, para luego exportar la información a un documento PDF mediante Apache FOP.

Dificultades

- Código en bash: si bien dentro de la carrera de Ingeniería Informática todos los integrantes del grupo vimos y fuimos evaluados en bash, el contenido limitado que vimos junto a la falta de práctica desde entonces provocó que el diseño del script fuera más complicado y tardara más de lo previsto.
- Atributos duplicados: durante la creación del archivo *nascar_data.xml* a partir de la consulta especificada, obtuvimos un error para ciertos parámetros donde un atributo de nombre “name” estaba duplicado. El problema era con los pilotos con más de un auto, pues estábamos

seleccionando todos los nombres de fabricantes de todos sus autos. Lo solucionamos eligiendo sólo el nombre del primer auto que se procesa.

- Ausencia del valor en el campo rank: notamos que si bien los pilotos estaban ordenados según su puntaje de la temporada, su rango no figuraba en el documento final. El problema era una condición, donde utilizamos un if que no cumplía con lo que necesitábamos y lo reemplazamos con un choose.

Aspectos investigados

- Programación en bash: como se mencionó en el punto anterior, debimos recurrir a notas y trabajos viejos para poder realizar el script necesario para automatizar los comandos de la consola y verificar la validez de los parámetros ingresados por el usuario.
- Creación y uso de la clave API: debimos entrar en el apartado de desarrolladores de la página web de SportRadar para obtener esta clave, que luego debimos utilizar en partes específicas del trabajo para recuperar la información de los pilotos.
- Creación del PDF: para poder pasar la información de la consulta a un PDF, debimos investigar sobre XSL-FO para poder generar un archivo que luego, tras instalar Apache FOP, podemos convertir a dicho formato.

Roles

- Responsable del funcionamiento de la consulta XQuery: Jonathan Lucas Blankleder.
- Responsable del funcionamiento de la plantilla XSLT: Jonathan Lucas Blankleder, Agustín Germán Ramírez Donoso.
- Responsable del funcionamiento global del proyecto: Ignacio Gastón Frund, Agustín Germán Ramírez Donoso.
- Responsable de la presentación: Ignacio Gastón Frund.

Conclusiones

Este trabajo requirió una buena administración de los tiempos de cada uno para finalizarlo a tiempo durante una época bastante ocupada, por lo que cada miembro del equipo debió realizar gran parte del trabajo en solitario antes de tener la oportunidad de que entre todos podamos ver y corregir los errores, y lograr que la totalidad del mismo funcione lo mejor posible. A pesar de ello fue interesante ver la aplicación de un caso real donde podemos acceder a un gran volumen de

información creado por otras personas y obtener lo que deseamos con una presentación que destaca los resultados más relevantes a nuestra búsqueda.