



Instituto Tecnológico de Buenos Aires

Trabajo Práctico Especial

Grupo 17

72.32 - Diseño y Procesamiento de Documentos XML
2do Cuatrimestre 2023

Autores:

Ana Negre	63074
Matías Leporini	62872
Mateo Cornejo	62722
Pablo Germano	62837

1. Introducción

Para el presente trabajo, se utilizaron las herramientas vistas a lo largo de la cursada para presentar y filtrar información de una base de datos de competencias de rugby en un formato adecuado. Para obtener estos datos se utilizó la **API** del sitio web SportRadar, junto a archivos **XQuery** para hacer las consultas pedidas, un **script de bash** para la ejecución de lo pedido manejando los errores y los datos extraídos y, por último, un **archivo XSLT** para llevar los mismos al formato Markdown, brindándole una estructura legible a la información. En este informe se documentan los procedimientos realizados, junto a las dificultades y aprendizajes encontrados durante la implementación de los objetivos propuestos.

2. Desarrollo

2.1 Trabajo realizado

Como hilo conductor de este trabajo, se creó el script bash *tpe.sh*. Éste recibe como parámetro el nombre de la temporada y el año en el que inició -ambos mediante la consola, al momento de ejecutar el script- y se encarga de entrelazar cada una de las partes.

En primer lugar, el script verifica que los parámetros sean los adecuados (i.e. que exista un parámetro de nombre y uno de año, este último mayor que 2007). En caso de que esta validación falle, el método *generate_error_xml* imprime un mensaje en *season_data.xml* para indicar el problema. Hecho eso, se extrae la información necesaria para *seasons.xml* con la key obtenida mediante la API de SportRadar.

Luego se hace la consulta indicada por *extract_season_id.xq*, devolviendo la misma el identificador de la primera temporada que coincida con los parámetros brindados (*season_id*). Utilizando la clave generada anteriormente y el identificador recién obtenido, se obtiene la información para *season_info.xml* (contiene *name*, *competition*, *date*, *year* y *stages*) y *season_lineups.xml* (contiene *name*, *type*, *date_of_birth*, *nationality* y *events_played*). A ambos se les remueve el Namespace para su correcto funcionamiento.

Con estos 2 XMLs, se ejecuta *extract_season_data.xq*. Esta consulta lleva a la generación del archivo *season_data.xml* (con año, nombre de la competencia, y el listado de sus corredores).

Finalmente, se crea el markdown mediante el XSLT *generate_markdown.xsl*. Este programa genera *season_page.md*, exhibiendo el mismo la información de *season_data.xml*, organizada de la forma indicada por la cátedra.

2.2 Dificultades encontradas

En el desarrollo del trabajo, surgieron las siguientes dificultades:

- Manejo de bash: Dado que ninguno de los miembros del equipo estaba familiarizado con bash, se requirió de un extenso proceso de prueba y error para la fabricación de *tpe.sh* (véase aspectos investigados).
- Creación de *generate_markdown.xsl*: Por motivos similares a bash, también fue necesario un proceso de prueba y error para comprender cómo crear un archivo markdown (nuevamente, véase aspectos investigados).
- Obtención de los jugadores en *extract_season_data.xq*: En las primeras iteraciones de este archivo, cada jugador aparecía repetido múltiples veces. Se requirió una revisión a fondo del código para resolver el problema.

2.3 Aspectos investigados

Para el trabajo, se investigaron los siguientes aspectos:

- Obtención de API key: Fue necesario averiguar cómo obtener la API de SportRadar.
- Validación de parámetros: Hizo falta la implementación de métodos de verificación para contemplar casos límite, como por ejemplo que no se pasen los parámetros correctos.
- Creación del markdown: Se buscó información adicional sobre la transformación XSLT del archivo *season_data.xml*, para lograr el resultado pedido.
- Ocultación header XML: Un aspecto que se tuvo que investigar fue la forma de elaborar una función en bash que permitiera borrar los namespaces de los archivos obtenidos de la API.
- Uso de Bash: Si bien se contaban con conocimientos básicos de bash, para el trabajo fue necesario revisar su uso, e investigar los comandos y sintaxis específicos para lograr la correcta ejecución del proyecto.

2.4 Roles de los integrantes

- Responsable del funcionamiento de la consulta xQuery: Ana Paula Negre
- Responsable de funcionamiento de la plantilla XSLT: Matías Leporini
- Responsable del funcionamiento del funcionamiento global del proyecto: Ana Paula Negre y Mateo Hernán Cornejo
- Responsable de la presentación: Matías Leporini y Pablo Omar Germano

3. Conclusiones

Este trabajo práctico concluye con una nueva noción de los distintos desafíos que pueden aparecer al implementar un programa. Múltiples problemáticas surgieron en su realización, pero cada una implicó un proceso de aprendizaje e investigación. Gracias a este proyecto, también comprendimos la importancia de las técnicas y prácticas vistas a lo largo de la cursada. El trabajo, en conclusión, nos permitió profundizar nuestro entendimiento sobre las mismas, además de mejorar nuestras habilidades de resolución de problemas mediante el manejo de datos. Esto, a su vez, nos permite estar mejor preparados para resolver futuros desafíos.