

Diseño y Procesamiento de Documentos XML

Segundo Parcial en Modalidad TPE

1er Cuatrimestre 2016

1. Objetivo

El objetivo de este trabajo práctico es aplicar las distintas herramientas de consulta y transformación de documentos XML vistas a lo largo del curso.

2. Contexto

Actualmente, los sistemas de Internet intercambian información por medio de documentos XML o, en una proporción cada vez mayor, mediante documentos de tipo JSON (JavaScript Object Notation), un estándar abierto basado en texto, entendible por humanos y pensado también para intercambio de datos.

Para facilitar la interoperabilidad entre sistemas, en ocasiones se requiere conectar interfaces que aceptan documentos XML con interfaces que aceptan documentos de tipo JSON.

Para lograr este intercambio de datos, se pueden desarrollar herramientas que automaticen la conversión de un documento XML a uno JSON, mediante la aplicación de las técnicas vistas en clase (XSLT, XPath, XQuery).

Para conocer la sintaxis de JSON se sugiere consultar:

<http://www.ecma-international.org/publications/standards/Ecma-404.htm>

http://www.w3schools.com/json/json_syntax.asp

3. Modalidad

El Trabajo Práctico deberá realizarse en **grupos**, conformados por **exactamente 3 personas**.

Los entregables del trabajo, compuestos por los ítems detallados en el punto 6, deben entregarse a través de la plataforma Sakai hasta el martes 21 de junio a las 08:00 AM.

Se evaluará el trabajo en presencia de los alumnos de cada grupo durante un intervalo de 30 minutos.

Los horarios estarán disponibles en el calendario de Sakai.

4. Tareas a realizar

El gobierno de la ciudad de Buenos Aires, pone a disposición pública conjuntos de datos con información referente a distintos eventos que se producen en el marco de ciertos festivales de la ciudad.

Dada una **fecha** y tomando como base los datos del festival de cine BAFICI 2013 que se encuentran en el archivo **bafici13-programacion.xml**, se deberá generar un documento JSON que contenga los eventos que se llevarán a cabo en el contexto del festival, en la fecha indicada.

5. Pasos a seguir

5.1. Bajar de Sakai el conjunto de archivos en formato xml:

Programación 2013: bafici13-programacion.xml,

Películas 2013: bafici13-films.xml,

Sedes 2013: bafici13-sedes.xml,

Países 2013: bafici13-paises.xml,

y el archivo **datos_jsonE.xsd**.

5.2.

Crear la consulta XQuery **junta.xq** que, tomando como base **bafici13-programacion.xml** y los IDs correspondientes en el resto de los .xml, produzca como salida el archivo **datos_json.xml**.

El formato final de **datos_json.xml** debe tener la estructura indicada en **datos_jsonE.xsd**.

La raíz del documento se llama **<peliculas>** y tiene un atributo fecha con la fecha que será pasada como parámetro.

Por cada película programada en la fecha indicada se creará un elemento **<pelicula>** cuyo contenido será:

Elemento	Contenido
<título>	El título de la película que figura en bafici13-films.xml
<descripción>	El contenido de <sinopsis_es> que figura en bafici13-films.xml
<lugar>	Tendrá tres subelementos <nombre><dirección><sala> explicados en la Tabla 2
<hora>	El contenido de <time_> en bafici13-programacion.xml
<nacionalidad>	Contiene tantos elementos <pais> como países haya en bafici13-films.xml correspondiente cuyo valor sea distinto de 0 y de NULL. Ver Tabla 3

Tabla 1 - Contenido de **<pelicula>**

Elemento	Contenido
<nombre>	Nombre de la sede como figura en bafici13-sedes.xml
<dirección>	Dirección de la sede como figura en bafici13-sedes.xml
<sala>	Nombre de la sala que figura en bafici13-programacion.xml

Tabla 2 - Subelementos de **<lugar>**

Elemento	Contenido
<pais>	Nombre del país tal como figura en <name_es> en bafici13-paises.xml

Tabla 3 - Subelementos de **<nacionalidad>**

En caso que el ID no exista, los alumnos tienen la libertad de elegir la acción a tomar la cual deberá ser documentada en el informe final.

Sólo se deben mostrar las películas: si es otro tipo de evento no debe aparecer en la salida.

5.3. Crear la plantilla de transformación **toJson.xslt** que tome los datos del archivo **datos_json.xml** creado en **5.2** y produzca como salida el archivo final **peliculas.json** con el formato y contenido siguientes:

Formato del JSON

El documento Json tendrá dos etiquetas principales: **“fecha”** y **“peliculas”**

- **“fecha”** tendrá el valor del atributo “fecha” del elemento **<peliculas>** de **datos_json.xml**.
- **“peliculas”** contendrá una lista de etiquetas **“pelicula”** con los datos de **<pelicula>** de **datos_json.xml**.

A continuación se muestra un ejemplo de cómo debería ser la estructura del documento JSON.

```
{ "fecha": "2013-04-11",
  "peliculas": [
    { "pelicula": {
      "titulo" : "tit1",
      "descripcion" : "des1",
      "lugar" : {
        "nombre" : "xxxx",
        "direccion" : "dir1",
        "sala" : "Sala1"
      },
      "hora" : "13:30:00",
      "nacionalidad" : [
        { "pais" : "pais1" },
        { "pais" : "pais2" }
      ]
    }
  },
  { "pelicula": {
    "titulo" : "",
    "descripcion" : "des2",
    "lugar" : {
      "nombre" : "yyyyy",
      "direccion" : "dir2",
      "sala" : "sala2"
    },
    "hora" : "11:00:00",
    "nacionalidad" : [
      { "pais" : "pais3" },
      { "pais" : "pais4" }
    ]
  }
}
]
```

5.5.

Todos los comandos a partir del punto **5.3** deben incluirse en el archivo **trabajo.sh** que recibirá como parámetro la fecha de la que se quieren mostrar las películas y servirá para la evaluación de los docentes.

5.6.

Para **5.2** y **5.3** **NO** se pueden usar librerías predefinidas. Se deben utilizar las herramientas vistas en clase: XSLT, XPath, XQuery.

6. Entregables

Los alumnos deberán entregar, a través de Sakai, sección **Tareas**, los siguientes documentos:

- **junta.xq**: la consulta XQuery explicada en **5.2**.
- **toJson.xsl**: la transformación explicada en **5.3**.
- **trabajo.sh**: el archivo con los comandos necesarios que permitan a los docentes, a partir de los documentos xml iniciales y una **fecha**, generar los archivos **datos_json.xml** y **peliculas.json**.
- Un **informe** documentando el trabajo realizado. En dicho informe, cuya extensión máxima no puede ser más de 3 carillas, se debe indicar:
 - El flujo del programa;
 - Las herramientas utilizadas en cada paso;
 - Las decisiones tomadas;
 - Lo que se haya tenido que investigar que no se haya dado en clase;
 - Y todo lo que considere necesario aclarar para su ejecución.

7. Evaluación

Los docentes ejecutarán **trabajo.sh** en el entorno Linux de los laboratorios.

Durante la evaluación los docentes probarán el trabajo usando distintas fechas, incluso fechas inexistentes para analizar el manejo de errores, y se validará el archivo **datos_json.xml** con el xsd esquema **datos_jsonE.xsd**.

Se evaluará fundamentalmente el trabajo en equipo y la participación de todos los miembros.

Asimismo, se realizarán preguntas a los alumnos para determinar los conocimientos de XPath, XQuery y XSLT aplicados a este trabajo.