

# Diseño y Procesamiento de Documentos XML

## Segundo Parcial en Modalidad TPE

### 2do Cuatrimestre 2015

#### 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

El gobierno de la ciudad de Buenos Aires, pone a disposición de los ciudadanos, datasets con información referente a distintos eventos que se producen en el marco de ciertos festivales. Estos datasets o conjuntos de datos, hacen referencias entre ellos que deben traducirse en un formato amigable para el acceso humano.

#### 3. Modalidad

El TP 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 18 de Noviembre 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 disponibles se encuentran en el calendario de Sakai

#### 4. Tareas a realizar

Dada una **fecha** y tomando como base los datos del festival DANZA CONTEMPORANEA 2012 que se encuentran en el archivo **fdc-eventos-2012.csv**, se deberá generar una página html que contenga una tabla con los eventos que se llevarán a cabo en el contexto del festival, en la fecha indicada.

#### 5. Pasos a seguir

5.1. Bajar de <http://data.buenosaires.gob.ar/dataset/ba-danza-contemporanea> el conjunto de archivos en formato csv (comma separated values):

**Eventos 2012:** fdc-eventos-2012.csv

**Obras 2012:** fdc\_obras-2012.csv

**Sedes 2012:** fdc-sedes-2012.csv

**Artistas 2012:** fdc-artistas-2012.csv

5.2. Convertir los archivos del punto 5.1 a formato xml conservando sus nombres originales.

Existen distintos mecanismos para convertir un archivo csv a xml. Por ejemplo, se lo puede hacer manualmente utilizando la página <http://codebeautify.org/csv-to-xml-json>.

Otra forma es utilizar la plantilla csv-to-xml\_v2.xslt disponible en Sakai.

Para usar dicha plantilla, si se quisiera transformar por ejemplo, **fdc-eventos-2012.csv** se debe correr la siguiente línea de comando:

```
Java -cp saxon9he.jar net.sf.saxon.Transform -o:fdc-eventos-2012.xml -it: main csv-to-xml_v2.xslt pathToCSV=fdc-eventos-2012.csv
```

Queda a criterio de los alumnos el mecanismo a utilizar, el cual debe indicarse detalladamente en el informe final.

**5.3.** Crear la consulta xquery **junta.xq** que, tomando como base **fdc-eventos-2012.xml** y las referencias indicadas en **5.4**, produzca el archivo **datos\_tabla.xml** con los datos que contendrá la tabla final pero en formato xml. En este paso, todas las referencias deberán ser reemplazadas por el contenido en el correspondiente archivo referenciado. Los datos de **datos\_tabla.xml** deberán sólo corresponder a la fecha indicada que será pasada como parámetro a la consulta.

El formato final de **datos\_tabla.xml** queda a criterio de los alumnos. Se recomienda crear un formato orientado a completar el punto **5.4**.

Por ejemplo, **datos\_tabla.xml**, podría tener un formato similar a

```
<tabla>
  <fila>
    <titulo>...</titulo>
    <descrip>..</descrip>
    <lugar>...</lugar>
    <hora>...</hora>
    <artista>...</artista>
    <bio>...</bio>
  </fila>
  <fila>
    ...
  </fila>
</tabla>
```

**5.4.** Crear la plantilla de transformación **eventos.xslt** que tome los datos del archivo **datos\_tabla.xml** creado en **5.3** y produzca como salida el archivo final **eventos.html** con el formato y contenido siguientes:

#### **Formato de la página html**

El **título** que aparecerá en la página será el nombre del festival

El **subtítulo** será: "Eventos del día: " (indicar la *fecha*)

La tabla debe contener el siguiente encabezado:

Título	Descripcion	Lugar	Hora	Artista	Biografía
--------	-------------	-------	------	---------	-----------

y tendrá una fila por cada obra que se lleve a cabo en la fecha indicada. Los datos de la tabla deben obtenerse de los distintos archivos relacionados ya que inicialmente sólo se cuenta con las referencias.

### Explicación de las columnas

**Título** es el nombre de la obra tal como figura en **shakespeare-obras-2012** (TITLE)

**Descripción** tendrá la sinopsis de la obra tal como aparece en SYNOPSIS\_ES del archivo **fdc\_obras-2012**.

**Lugar** estará formado por el **nombre** del lugar donde se desarrolla la obra y su **dirección**. Son los campos TITLE y ADDRESS de **fdc-sedes-2012**

**Hora** es la hora en que se llevará a cabo el evento en la fecha indicada, (TIME)

**Artista** tendrá el nombre del primero de los artistas, referenciado por ID\_ARTIST1 tal como aparecen en el archivo **fdc-artistas-2012**. (NAME)

**Biografía** tendrá la biografía del artista referenciado por ID\_ARTIST1 tal como aparece en el archivo **fdc-artistas-2012** (BIO\_ES)

En estos últimos casos, el ID del artista debe tomarse de **fdc\_obras-2012.xml**.

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.

Optativamente, se pueden agregar hasta dos columnas más que enriquezcan la información de la tabla cuya explicación debe ser incluida en el informe final.

**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 los eventos y servirá para la evaluación de los docentes

### 6. Entregables

Los alumnos deberán entregar los siguientes documentos

- **fdc-eventos-2012.xml, fdc\_obras-2012.xml, fdc-sedes-2012.xml, fdc-artistas-2012.xml**
- **junta.xq**: la consulta XQuery explicada en **5.3**.
- **eventos.xsl**: la transformación explicada en **5.4**
- **trabajo.sh**: el archivo con los comandos necesarios que permitan a los docentes, a partir de los documentos xml producidos en **5.2** y una **fecha**, generar los archivos **datos\_tabla.xml** y **eventos.html**
- Un informe documentando el trabajo realizado.

### 7. Evaluación

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

Durante la evaluación los docentes probarán el trabajo usando distintas fechas inclusive, fechas inexistentes para analizar el manejo de errores.

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, XSLT aplicados a este trabajo.