



Procesamiento de información en aplicaciones telemáticas Curso 2023/2024

Práctica 1

XMLSchema

Versión 8/marzo/2024

Tabla de contenido

1	Objetivos.....	1
2	Presentación	1
3	Descripción del modelo	1
3.1	Documento de identificación de los responsables de una misión	1
3.2	Documento de descripción de un robot.....	2
3.3	Documento de descripción de una tarea	3
3.4	Documento de descripción de la misión	4
4	Modelos de información.....	5
5	Realización.....	6
5.1	FASE 1.....	6
5.2	FASE 2.....	7
5.3	FASE 3.....	7
6	Entrega	8
7	Referencias	9

1 Objetivos

Esta práctica tiene como objetivo aplicar los conocimientos adquiridos en las clases de teoría sobre:

- Generación de documentos XML bien formados.
- Definición de esquemas de documento mediante *XMLSchema*.
- Generación de documentos XML válidos con respecto a un esquema de documento.

2 Presentación

Esta primera práctica de la asignatura le servirá para ejercitar los conocimientos adquiridos en las sesiones de teoría correspondientes al tema 2: parte 1 y 2.

Dispone de tres semanas para realizarla, teniendo en cada semana una sesión presencial de laboratorio. En la semana siguiente a la última sesión, se realizará la evaluación de la práctica dentro del examen del primer parcial de la asignatura.

Como puede ver en el apartado [Realización](#), la práctica debe desarrollarla en 3 fases, una por semana. En cada fase generará un conjunto de ficheros que debe guardar en un directorio diferente. Al final, tal y como dice el apartado [Entrega](#), deberá crear un fichero comprimido que contenga los tres directorios (uno por cada fase) con los ficheros correspondientes.

Para trabajar con los documentos XML se le propone que use el programa *Visual Studio Code*. En el apartado [Referencias](#) de este documento encontrará información de un recurso de Moodle [1] que puede usar para instalar y configurar el programa con las extensiones adecuadas que le permitirán editar documentos XML y validarlos frente a un *XMLSchema*. De forma alternativa, también puede usar el programa *Eclipse* con los *plugins* adecuados, tal y como se explica en otro de los recursos [2] publicado en Moodle. No obstante, la recomendación es que use el programa *Visual Studio Code*, pues éste será el programa sobre el que los profesores de la asignatura darán soporte.

Es muy recomendable que asista a las sesiones presenciales de laboratorio, pues el profesor, además de atender dudas, le orientará en algunos aspectos que le facilitarán la realización de la práctica.

Esta práctica la debe realizar de forma individual.

3 Descripción del modelo

Una empresa de desinfección e higienización de lugares públicos ha decidido utilizar un conjunto de robots que permitan realizar las misiones de higienización de un espacio de forma segura para el ser humano. Una misión queda especificada mediante la definición de un conjunto de tareas. Cada una de ellas será ejecutada por uno o varios robots, de manera coordinada, implicados en la misma misión. Existe un centro de control en el que hay asignados uno o varios responsables para cada misión.

Se ha optado por utilizar, como soporte para la configuración de una misión, una serie de documentos XML que se describen en los siguientes apartados.

3.1 Documento de identificación de los responsables de una misión

El elemento raíz de este documento se denominará *responsable* y representará a una empresa que puede destinar a varias personas y/o departamentos al control de la misión. Tiene un atributo obligatorio, llamado *identificadorFiscal*, cuyo valor es el número de identificación ante el Ministerio de Industria de la empresa. Está formado por una letra mayúscula, seguida de siete caracteres numéricos y otra letra mayúscula.

El elemento *responsable* está formado por los siguientes elementos obligatorios (sin importar el orden de aparición):

- *colaboradores*: especifica las personas y/o departamentos de la empresa encargados de la supervisión de la misión. Es una lista no vacía de elementos, que pueden ser:
 - *persona*: cadena de caracteres con el nombre y apellidos de una persona.
 - *departamento*: cadena de caracteres que indica el nombre de un departamento.Puede haber varias personas y varios departamentos, por ejemplo, podría haber 3 elementos persona y 2 elementos departamento, o un solo elemento persona, o un solo elemento departamento.
- *contacto*: formado por un elemento *email* y un elemento *uri*, sin importar el orden de aparición:
 - *email*: dirección de correo electrónico. Obligatorio.
 - *uri*: página web. Opcional.

3.2 Documento de descripción de un robot

El elemento raíz de este documento se denominará *robot* y tendrá un atributo obligatorio de nombre *idRobot* con el número de identificación del robot. El valor del atributo estará formado por los caracteres *id* seguidos de un guion y tres caracteres numéricos. Ejemplo: "id-251".

Este elemento raíz se compone de una secuencia de los siguientes elementos, todos obligatorios excepto los que se indiquen expresamente como opcionales:

- *funcionalidad*: cadena de caracteres de longitud máxima 256 que explica, de manera resumida, la funcionalidad del robot.
- *sistemaLimpieza*: palabra que identifica el sistema de limpieza que utiliza el robot. Puede tomar uno de los siguientes valores: *ozono*, *vapor* o *ultravioleta*.
- *homologaciones*: lista de hasta diez palabras que especifican las homologaciones que ha pasado el robot. Este elemento es opcional.
- *sensores*: compuesto por la secuencia de los siguientes elementos:
 - *númeroSensores*: valor entero. Indica el número de sensores que tiene incorporados el robot. Este elemento es obligatorio.
 - Una secuencia no vacía de elementos *sensor*. Cada elemento *sensor* tiene un atributo obligatorio de nombre *idSensor* cuyo valor es una palabra formada por dos letras minúsculas seguidas de un carácter numérico y de dos letras mayúsculas que identifica el sensor. Ejemplo: "ox1AV".

El atributo *idSensor* debe tomar un valor único dentro del conjunto de sensores del robot.

Cada elemento *sensor* se compone de una secuencia de los siguientes elementos, todos obligatorios:

- ✕ *patógeno*. Una palabra que especifica el patógeno que detecta el sensor. Ejemplo: "estafilococo".

- ✧ *densidad*. Una cadena de caracteres formada por un número seguido de una palabra que indica la unidad en la que se especifica la densidad. Ejemplo: "5000 bacterias/m2".
- *actuadores*: compuesto por la secuencia de los siguientes elementos:
 - *númeroActuadores*: valor entero. Indica el número de actuadores que tiene instalados el robot. Este elemento es obligatorio.
 - Una secuencia no vacía de elementos *actuador*. Cada elemento *actuador* tiene un atributo obligatorio de nombre *idActuador* cuyo valor es una palabra formada por dos letras mayúsculas seguidas de dos caracteres numéricos que identifica el actuador. Ejemplo: "AC01".

El atributo *idActuador* debe tomar un valor único dentro del conjunto de actuadores del robot.

Cada elemento *actuador* se compone de una secuencia de los siguientes elementos, todos obligatorios:

- ✧ *descripción*. Una cadena de caracteres que describe las funciones del actuador.
- ✧ *alcanceActuación* compuesto por los siguientes elementos obligatorios:
 - *tipoActuación*: un texto que especifica la actuación.
 - *fechasActuación*: una lista no vacía que puede contener valores de tipo *time* o *datetime* indistintamente. Un valor de tipo *datetime* indica que la actuación se realiza ese día a esa hora y un valor de tipo *time* indica que la actuación se realiza de forma periódica todos los días a la hora indicada. Esta lista no puede tener más de 10 valores.

3.3 Documento de descripción de una tarea

El elemento raíz de este documento se denominará *tarea* y tendrá un atributo obligatorio de nombre *idTarea* formado por una secuencia de cinco letras mayúsculas seguidos de un guion y 4 dígitos numéricos. Ejemplo: "TOXXI-0001"

Este elemento raíz se compone de una secuencia con los siguientes elementos, todos obligatorios:

- *descripciónTarea*: una cadena de caracteres que describe la tarea.
- *espacioTarea*: determina el espacio en el que deberá ejecutarse la tarea. Puede ser uno (y solo uno) cualquiera de los siguientes elementos:
 - *volumenEspacio*. Describe un espacio de base poligonal en el que se va a realizar la tarea a partir de los siguientes elementos, sin importar su orden de aparición, pero siendo todos obligatorios:
 - ✧ *centro*: punto geográfico central de la base del espacio en la que se localiza la tarea a realizar. Consta de los siguientes elementos, que pueden aparecer en cualquier orden:
 - *longitud*: valor real entre 0 y 180.
 - *latitud*: valor real entre -90 y 90.
 - *altitud*: valor real entre -416,5 y 8848.

- ✖ *forma*: indica la forma que tiene la base del espacio en el que se va a realizar la tarea. Los valores posibles son: círculo, cuadrado, rectángulo o polígono.
- ✖ *altura*: expresa en centímetros, mediante un valor entero, la altura del espacio en el que se va a realizar la tarea.
- *puntosEspacio*. Sirve para describir un espacio con base poligonal irregular. Es una secuencia con los siguientes elementos obligatorios:
 - ✖ Una secuencia de al menos tres elementos *punto*. Cada elemento punto estará caracterizado por *longitud*, *latitud* y *altitud*, ajustándose estos elementos a la especificación dada para los elementos *centro*.
 - ✖ *altura*: valor entero. Expresa en centímetros la altura del espacio en el que se va a realizar la tarea.
- *imagen*. Se utiliza para representar una imagen. Este elemento tiene un atributo obligatorio llamado *idImagen* cuyo valor está formado por siete caracteres, de los cuales los cuatro primeros son IMAG y los tres últimos son numéricos. Ejemplo: "IMAG001". Este elemento contiene la secuencia de los siguientes elementos obligatorios:
 - ✖ *datosImagen*: formado por la siguiente secuencia de elementos:
 - *dimensiones*: elemento obligatorio compuesto por la secuencia de los siguientes elementos, también obligatorios, que sirven para representar en centímetros el ancho y el alto de la imagen mediante números reales:
 - *ancho*.
 - *alto*.
 - *descripción*: elemento opcional que contiene una cadena de caracteres de longitud indefinida.
 - ✖ *contenido* de la imagen. Puesto que el contenido de una imagen es un valor en binario, y en un documento XML solo puede haber texto, habrá que realizar una codificación a Base64Binary del binario de la imagen. No es relevante para esta práctica realizar esto, por lo que puede contener simplemente un texto¹ descriptivo. Este elemento tendrá un atributo, llamado *formato*, cuyo contenido se modelará mediante el uso de notaciones. Podrá ser GIF (image/gif), JPEG (image/jpeg) o PNG (image/png).

3.4 Documento de descripción de la misión

En este documento XML se va a reflejar toda la información de una misión concreta. Para ello se van a utilizar los elementos que pertenecen a los documentos de descripción del *robot*, de descripción de la *tarea* y de identificación de los *responsables* de la misión. En el documento se debe especificar el responsable o responsables de la tarea, el número y secuencia de actividades de cada tarea y el conjunto de robots encargados de llevar a cabo cada una de las tareas.

¹ Si desea poner el contenido de una imagen, puede realizar la codificación a Base64binary usando una herramienta de conversión de binario a texto on-line, como por ejemplo [ésta](#).

El elemento raíz del documento es *misión* y tendrá un atributo obligatorio de nombre *idMisión* que sirve como identificador de la misión. El valor del atributo estará formado por los caracteres MIS seguidos de tres caracteres numéricos. Ejemplo: "MIS001".

Este elemento raíz se compone de una secuencia de los siguientes elementos, todos obligatorios:

- *fecha*. Indica la fecha de ejecución de la misión en formato yyyy-mm-dd.
- *idioma*: elemento vacío que contendrá un atributo de nombre *lang* con el código de idioma que se especifica en la norma ISO 639-1.
- *descripción*: compuesto por una combinación de cadenas de caracteres y palabras clave (elementos *key* que identificarán conceptos clave dentro de la *descripción*). En una *descripción* podrán aparecer tantas palabras clave como sea necesario. Por ejemplo:

```
Misión: <key> higienizar </key> los <key> quirófanos </key> existentes  
en el <key> Hospital Universitario </key> para conseguir <key> eliminar  
bacterias </key>
```
- *responsables*: especifica los responsables de la misión. Contendrá uno o varios elementos *responsable* como los descritos en el apartado 3.1. Es necesario asegurar que el *identificadorFiscal* de cada responsable tenga un valor único en el documento.
- *robots*: contendrá uno o varios elementos *robot* como los descritos en el apartado 3.1. No puede haber dos robots con el mismo *idRobot* en el documento.
- *tareas*: contendrá uno o más elementos *tarea* como los descritos en el apartado 3.3. No puede haber dos tareas con el mismo *idTarea* en el documento.
- Una lista no vacía de elementos *robotTarea* cada uno de ellos formado por la secuencia de los siguientes elementos:
 - *tareaRealizada*: identifica la tarea mediante su *idTarea*. Ese identificador de corresponderá a una de las tareas presentes en el documento.
 - Entre 1 y 10 elementos *robotAsignado* que identifican a los robots que han llevado a cabo esa tarea mediante sus *idRobot*. Estos identificadores deben corresponder a robots presentes en el documento.

4 Modelos de información.

Para que el sistema gane en robustez es necesario que se establezcan los modelos de datos (esquemas) correspondientes a los documentos que se han realizado en el apartado anterior. Para ello se ha decidido utilizar el lenguaje de esquema *XMLSchema*.

Para desarrollar estos documentos esquema se han de tener en cuenta las siguientes consideraciones:

- Se ha de realizar un *XMLSchema* por cada uno de los documentos XML especificados en el apartado 3.
- Cada uno deberá tener su propio *espacio de nombres*, siendo este a su elección.
- El prefijo que se usará para los elementos del espacio de nombres de *XMLSchema* será *xsd*.

- No es necesario comprobar que el número de sensores y el número de actuadores de un robot coincide con los valores reflejados en *númeroSensores* y *númeroActuadores* respectivamente.
- En todos los elementos `<xsd:schema>` deberá figurar el atributo `elementFormDefault = "qualified"`.
- Para indicar el tipo de los elementos que identifican la *latitud* y *longitud* en el esquema de *tarea* debe usar un esquema de amplia difusión como puede ser KML, GML, GPX, Este esquema se deberá incorporar al esquema de *tarea* y se tendrán en cuenta en el documento de descripción de la tarea para esos elementos.

5 Realización

El objetivo final de la práctica es realizar la especificación de datos completa. Para conseguirlo deberá ir haciendo el trabajo de forma incremental. Se le propone que lo haga en tres fases, donde en cada una de ellas deberá generar los documentos XML que se le indican. Para ello cree en su espacio de trabajo una carpeta por cada fase y en ella escriba los documentos que se le piden a continuación.

Todos los ficheros deben usar la codificación UTF-8 y reflejarlo así en la declaración XML del prólogo del documento.

Debe poner al final del prólogo de todos los ficheros, como comentario XML, su nombre completo y su correo electrónico UPM.

5.1 FASE 1

Escriba los siguientes documentos *XML* bien formados:

- **Descripción de los responsables.** Debe realizar dos documentos que describan los responsables según las especificaciones del apartado 3.1. Los documentos que debe realizar son:
 - **Responsable1.xml:** debe contener 3 personas y 2 departamentos. También usará el elemento *uri*.
 - **Responsable2.xml:** debe contener un departamento y no debe usar el elemento *uri*.
- **Descripción de los robots.** Debe realizar dos documentos que describan los robots según las especificaciones del apartado 3.2. En cada robot use un valor diferente para el elemento *sistemaLimpieza*. Los documentos que debe realizar son:
 - **Robot1.xml:** Deberá contener al menos dos elementos *sensor* y un elemento *actuador*. Además, debe usar todos los elementos definidos como opcionales.
 - **Robot2.xml:** Deberá contener un elemento *sensor* y al menos dos elementos *actuador*. Además, no debe usar los elementos definidos como opcionales.
- **Descripción de las tareas.** Debe realizar tres ejemplos de descripción de una tarea según las especificaciones del apartado 3.3. Los documentos difieren en cuanto al elemento usado en *espacioTarea* siendo:
 - **Tarea1.xml:** debe usar el elemento *volumenEspacio*.
 - **Tarea2.xml:** debe usar el elemento *puntosEspacio*.

- **Tarea3.xml**: debe usar el elemento *imagen* y todos los elementos que lo componen, aunque sean opcionales.
- **Descripción de una misión**. Debe realizar un documento, de nombre **Mision.xml**, con la descripción de una misión concreta según las especificaciones del apartado [3.4](#). La misión estará compuesta por dos elementos *robot*, tres elementos *tarea*, dos elementos *responsable* y al menos dos elementos *robotTarea* para la asignación de los robots definidos a una o varias de las tareas que componen la misión. Debe incluir los ficheros XML correspondientes a los responsables, los robots y las tareas usando **XInclude**. De esta forma se aprovecha el diseño modular.

La información contenida en los documentos no tiene que ser real, pero tiene que cumplir los requisitos especificados en la descripción del modelo.

5.2 FASE 2

Escriba los documentos *XMLSchema* que describen la estructura y las restricciones de los contenidos de los documentos XML especificados en el apartado [3](#). Los documentos que debe realizar son:

- **Responsable.xsd**: describirá el esquema de los *responsables* según las especificaciones del apartado [3.1](#). Para especificar el contenido del elemento *email* no es necesario que use una expresión regular compleja para asegurar que el valor tenga el formato de un correo electrónico válido. Es suficiente con restringir los valores a una cadena de caracteres que contenga en medio el carácter @. En la parte 3 del tema 2 aprenderá a realizar expresiones regulares complejas como la que se requeriría para este elemento.
- **Robot.xsd**: describirá el esquema de los *robots* según las especificaciones del apartado [3.2](#).
- **Tarea.xsd**: describirá el esquema de las *tareas* según las especificaciones del apartado [3.3](#).

Para realizar estos documentos debe tener en cuenta las consideraciones realizadas en el apartado [4](#).

Además de los documentos *XMLSchema* anteriores, deberá copiar los documentos XML de la fase 1 correspondientes a los representantes, robots y tareas, a los que le añadirá, en el elemento raíz, el atributo `schemaLocation` para indicar el documento *XMLSchema* con el que se tiene que validar. De esta forma podrá verificar si ha cometido algún error en los documentos *XMLSchema* o en los XML.

En todos los documentos (*XML* y *XMLSchema*) se deberán prefijar todos los espacios de nombres que se utilicen.

5.3 FASE 3

El objetivo de esta fase es realizar el documento *XMLSchema* que describa la estructura y las restricciones del contenido de un documento XML instancia de una misión. Para ello deberá realizar los siguientes documentos:

- **Mision.xsd**: describirá el esquema de las *misiones* según las especificaciones del apartado [3.4](#).

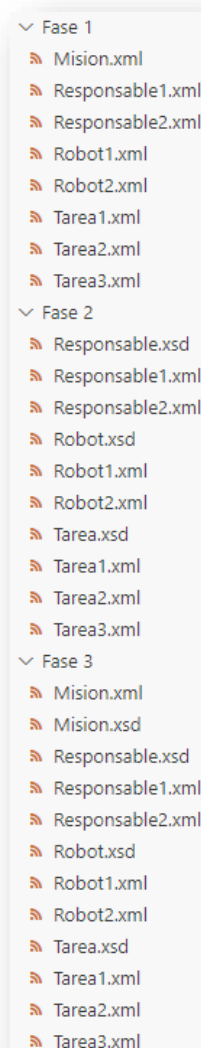
- **Mision.xml**: instancia de una misión, compuesto por dos elementos *robot*, tres elementos *tarea*, dos elementos *responsable* y los elementos *robotTarea* que considere convenientes para la asignación de los robots definidos a una o varias de las tareas que componen la misión.

Para poder realizar esta fase tenga en cuenta los siguientes aspectos:

- Debe copiar los ficheros utilizados en la fase anterior.
- Debe copiar el fichero Mision.xml de la Fase 1 y adaptarlo para que se pueda validar contra el esquema definido en el fichero Mision.xsd.
- El esquema de la *misión* usará los otros tres esquemas (*robot*, *tarea* y *responsable*).
- Debe incluir en el fichero Mision.xsd las restricciones especificadas en el apartado [3.4](#).

6 Entrega

Con anterioridad al **11 de marzo de 2024 a las 15:30 horas**, deberá entregar en Moodle, dentro del recurso “Espacio para la entrega de la práctica 1”, un fichero comprimido que incluirá los ficheros *.xml* y *.xsd* correspondientes a las tres fases de la práctica. En la siguiente imagen puede ver la estructura de directorios y ficheros que deberá contener el fichero comprimido que entregue:



Recuerde que debe poner al principio de todos los ficheros, como comentario XML, su nombre completo y su correo electrónico UPM.

Es obligatorio entregar esta práctica para poder presentarse al examen del primer parcial del día 12 de marzo.

7 Referencias

Disponible en la plataforma Moodle de la asignatura:

-
- [1]. [Instalación y configuración de Visual Studio Code](#)
 - [2]. [Instalación de plugins XML en eclipse](#).
-