

# SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN EN LA WEB

## TALLER No. 1

## XML & XSD

**Profesor:** Jaime Alberto Guzmán Luna

Contenido del taller:

1. XML
2. XSD

Primero, para el desarrollo del taller se recomienda el uso de un editor de texto como **Notepad++** o el recomendado para la clase **VSCode**.

Se recomienda a su vez la instalación de la extensión XML(XML Language Support)



Y al momento de estar dentro de un archivo XML puede presionar las teclas ctrl + shift + m para desplegar el apartado de problemas y revisar los errores que tenga el documento XML.

### ¿Qué se entrega?

Entregue los archivos XML y XSD y en un archivo txt, word, pdf, o el de su preferencia, las respuestas a las preguntas según sea el caso.

### EJERCICIO 1

Transcriba lo siguiente, a un archivo XML, e identifique algunos de los elementos mas relevantes que se observan en la misma.

# SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN EN LA WEB

```
<?xml version="1.0" encoding="UTF-8"?>
<!--SISTEMAS DE RECUPERACION WEB-->
<libros xmlns="https://books.org/">
  <libro ISBN="1613820917">
    <titulo>Hamlet</titulo>
    <autor>W. Shakespeare</autor>
    <paginas>330</paginas>
    <editorial>Simon & Brown</editorial>
  </libro>
  <libro ISBN="9789584254726">
    <titulo>La sombra del viento</titulo>
    <autor>Carlos Ruiz Zafon</autor>
    <paginas>575</paginas>
    <editorial>Planeta Booket</editorial>
  </libro>
  <libro ISBN="0765367297">
    <titulo>&lt; Halo &gt;; The Fall of Reach</titulo>
    <autor>Eric Nylund</autor>
    <paginas>448</paginas>
    <editorial>Tor Books</editorial>
  </libro>
</libros>
```

## EJERCICIO 2

Haciendo uso de la extensión o del siguiente validador WEB para XML

<https://jsonformatter.org/xml-editor> . Identifique los errores del siguiente código XML e indique como corregirlos.

```
<?xml version="\1.0">
<--Sistemas de Recuperacion-->
<mensaje>
<remite.>
nombre id=8976>Alfredo Reino</nombre>
<email>alf@ibium.com</email>
</remite>
<destinatario>>
<nommbre>>
<nombre id=0976>Bill Clinton</nombre>
<email>>president@whitehouse.gov
</destinatario/>
<asunto">Hola Bill</asunto>
<texto.>
  <[CDATA[
    <mensaje> trabajo de clase <>
      ejemplos
  ]
  <<parrafo>¿Hola qué tal? Hace <enfasis->mucho</enfasis> que no escribes. A ver si llamas
  y quedamos para tomar algo.</parrafo>
</texto>
</mensaje>
</mensaje>
```

## EJERCICIO 3

# SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN EN LA WEB

Represente 5 productos de **LEGO** y algunas de sus propiedades usando XML, represente 3 propiedades como atributos y 3 propiedades como elementos use el siguiente enlace para observar los productos y propiedades: <https://www.target.com/c/lego-technic/-/N-k5gdj>

## Espacios de Nombre (Namespace) en XML

Es muy posible que durante el desarrollo de un documento XML, nos encontremos frente a la dificultad de poder diferenciar dos elementos que se llamen igual pero que pertenecen a contextos totalmente diferentes, por ejemplo, diferenciar entre el peso (propiedad de un cuerpo), el peso (moneda colombiana) y el peso (moneda argentina). Para solucionar este inconveniente tenemos a nuestra disposición los espacios de nombre. Los espacios de nombre pueden ser declarados atributos del elemento raíz o como atributo del elemento que hará uso del mismo.

La sintaxis para declarar un namespace como atributo al interior de un elemento es la siguiente: `xmlns:prefijo="URI"`

De acuerdo con lo anterior, el problema propuesto podría solucionarse de la siguiente forma:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<Pesos xmlns:col="http://es.wikipedia.org/wiki/Peso_colombiano" xmlns:arg="http://es.wikipedia.org/wiki/Peso_de_Argentina">
  <col:Peso>
    <!-- Valor en dólares -->
    <col:Valor>0.000564</col:Valor>
  </col:Peso>
  <Peso>Propiedad Física de los cuerpos.</Peso>
  <arg:Peso>
    <arg:Valor>0.230840</arg:Valor>
  </arg:Peso>
</Pesos>
```

**Nota:** El URI del namespace no es utilizado por el analizador para buscar información, es usado únicamente para darle unicidad al namespace. Sin embargo, generalmente se utiliza como URI una dirección que contenga información referente al namespace.

## EJERCICIO 4

Tome el siguiente documento XML

```
<?xml version="1.0" encoding="UTF-8"?>
<mc:micasa
  xmlns:mc='http://www.geneura.org/micasa'
  xmlns:mueble='http://www.geneura.org/mueble'>
</mc:micasa>
```

Haga uso de este como base para construir un documento que tenga las características siguientes haciendo uso de los espacios de nombre en vez de sus direcciones de donde provienen los elementos:

1. mc:micasa contendrá un elemento habitación proveniente del vocabulario 'http://www.geneura.org/micasa' y que contiene el atributo id="comedor".

# SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN EN LA WEB

2. El anterior elemento habitación contendrá un elemento mueble del vocabulario 'http://www.geneura.org/micasa' elemento quien a su vez contendrá:
- un elemento nombre del vocabulario 'http://www.geneura.org/mueble' quien contiene una cadena de caracteres Sofá
  - Un elemento descripción del vocabulario 'http://www.geneura.org/mueble' quien contiene una cadena de caracteres Peludo

## EJERCICIO 5

Cree un archivo XML schema (XSD) que defina una familia, la familia está compuesta por un conjunto de únicamente de 4 personas, cada persona debe contener un nombre, un apellido, una edad entre 0 y 120, al menos definir 2 hijos, incluyendo el nombre o la edad de los hijos de esa persona, y representar el mes de nacimiento, estas propiedades deben definirse en este orden.

**Nota:** Tenga en cuenta los tipos de datos, y en la propiedad mes usar restricciones para definir los meses del año.

Use [http://www.utilities-online.info/xsdvalidation/#.U-AEI\\_mBN8E](http://www.utilities-online.info/xsdvalidation/#.U-AEI_mBN8E) para validar los XSD o la extensión de VSCode, esta se encarga de verificar sintaxis.

## EJERCICIO 6

Realizar un documento XML que tenga como modelo de datos el Esquema XML realizado en el ejercicio 5 y válidelo con la extensión de VSCode o el enlace del anterior ejercicio.

**Nota:** para referirse a la ubicación del XSD al momento de definir un XML puede hacerlo usando `<?xml-model href="ubicacion"?>` en la primer línea del archivo XML.

## MATERIAL DE APOYO

- Tutorial XML: <https://www.w3schools.com/xml/default.asp>
- Tutorial XSD: [https://www.w3schools.com/xml/schema\\_intro.asp](https://www.w3schools.com/xml/schema_intro.asp)