



PRÁCTICA Nº 3: 3 sesiones

INTRODUCCIÓN A XML Y DTD

0. OBJETIVOS

El objetivo de esta práctica es iniciarse en el lenguaje de marcas XML, en su sintaxis particular y poniendo especial atención en la elaboración de las etiquetas y el marcado de los documentos.

Para llegar a estos objetivos, entraremos en contacto con las DTD (definición de tipo de documento) y a partir de su elaboración marcaremos documentos.

A partir de las DTD y los documentos marcados XML llegaremos a la comprensión del llamado documento **bien formado** y del documento **válido**.

1. INTRODUCCIÓN

Para el desarrollo de la práctica utilizaremos un editor de ficheros XML y DTD, Cooktop 2.5. El editor es gratuito, por lo que puede ser descargado de la página web de la asignatura para ser instalado en nuestro ordenador en casa.

Con este tipo de herramientas podemos de forma fácil editar las DTD, y a partir de estas especificaciones construir nuevos ficheros XML o marcar ficheros electrónicos y validarlos.

Los ficheros son fáciles de seguir ya que se marcan en diferentes colores los diferentes elementos que componen los ficheros.

2. REALIZACIÓN DE LA PRÁCTICA

Se ofrece al alumno un archivo de texto, DSI.TXT. Este archivo contiene el supuesto envío periódico de información de un sistema de distribución selectiva de la información (DSI). Se supone que un usuario se ha suscrito a un servicio de este tipo para que le notifiquen las publicaciones relativas a un tema específico de forma periódica que van apareciendo en una base de datos, es decir, cada cierto tiempo, el sistema realiza automáticamente una búsqueda en una base de datos, en este caso *Information Science Abstracts*, con unas palabras clave elegidas por el usuario, en este caso *bibliometrics* y *science*; tras realizar la búsqueda, se le envía por correo electrónico un mensaje con un fichero como el que se provee al alumno.

Si observamos una parte del fichero:

```
Distribución Selectiva de la Información: bibliometrics, science
14-04-2005
La búsqueda: «bibliometrics» en la base de datos Information Science &
Technology Abstracts Plus 1966-2004/12 obtuvo 10 registros

Registro 1 de 10 en Information Science & Technology Abstracts Plus 1966-2004/12

TITLE: A review of the development and application of the Web impact factor.
SOURCE (BIBLIOGRAPHIC CITATION): Online-Information-Review. November 26, 2003;
27(6): 407-417
ISSN PRINT VERSION: 1468-4527
JOURNAL URL: http://www.emeraldinsight.com
```

PUBLICATION YEAR: 2003

LANGUAGE: English

ABSTRACT: Since 1996, hyperlinks have been studied extensively by applying existing bibliometric methods. The Web impact factor (WIF), for example, is the online counterpart of the journal impact factor. This paper reviews how this link-based metric has been developed, enhanced and applied. Not only has the metric itself undergone improvement but also the relevant data collection techniques have been enhanced. WIFs have also been validated by significant correlations with traditional research measures. Bibliometric techniques have been further applied to the Web and patterns that might have otherwise been ignored have been found from hyperlinks. This paper concludes with some suggestions for future research.

DESCRIPTORS: BIBLIOGRAPHY-; BIBLIOMETRICS-; INFORMATION-services; RESEARCH-

ACCESSION NUMBER: 3902810

Puede observarse que el envío que realiza el sistema DSI consiste en una descripción general de la DSI con su fecha de envío y las características de la búsqueda solicitada. A continuación, se ofrece una sucesión de varias referencias bibliográficas correspondientes a los registros generados por la búsqueda. Para cada una de estas referencias, se han obtenido datos como la fuente, los autores, el título, el abstract y demás.

Los objetivos de la práctica son:

1. Analizar la estructura del documento DSI.TXT.
2. Crear una DTD que describa la estructura de los documentos generados por la DSI y que sirva para marcar en XML estos textos que recibe el usuario. Esta DTD debe describir la estructura hasta el máximo nivel de detalle.
3. Generar un documento XML, resultante de marcar el documento DSI.TXT según la DTD definida en el punto anterior.

NOTA IMPORTANTE: En el fichero DSI.TXT se han dejado los nombres de campo tales como AUTHOR o LANGUAGE sólo para que sirvan de orientación al alumno en el análisis del documento. En el fichero XML creado por el alumno **NO** deberán aparecer estos nombres de campo, ya que dicho fichero estará marcado por las etiquetas que el alumno haya definido en su DTD.

4. Comprobar que el documento XML está bien formado y es válido respecto a la DTD definida.

3. ENTREGA DE MEMORIAS

Las memorias de las prácticas se entregarán por correo electrónico **desde la cuenta de correo electrónico de la Universidad de Valencia** al profesor asignado antes del día del examen. En el mensaje deberán incluirse los nombres, apellidos y DNI de los integrantes del equipo de prácticas, así como los dos ficheros desarrollados durante las prácticas: DSI.dtd, DSI.xml.

El profesor, tras comprobar el documento, enviará un mensaje de confirmación en las cuarenta y ocho horas siguientes. Si el alumno no recibiese dicho mensaje de confirmación en el plazo citado deberá ponerse en contacto con el profesor.

En ningún caso se aceptarán memorias por cualquier otro medio que no sea el correo electrónico.

PLAZO DE ENTREGA DE MEMORIAS: Una semana después de la tercera sesión de esta práctica.