



## XML: un nuevo lenguaje, una nueva necesidad

En el decenio en curso, la edición de las revistas médicas ha dado un vuelco espectacular hacia una nueva era: la de la información digital. Hoy día, no hay escuela de medicina, hospital, sociedad, agrupación o colegio médico que no requiera tener su propia revista. Y como esto sucede en todo el mundo, la cantidad de revistas médicas es cercana a los 45 mil títulos. Y las hay desde semanarios, quincenales, mensuales, bimestrales y con periodicidades más extendidas. Distribuir las impresas en papel es impensable y un atentado al medio ambiente, pues se requeriría la tala de millones de árboles para producir suficiente papel. Pensar en colocarlas en estantes de grandes bibliotecas para que puedan consultarse también es ahora inviable, porque no habría tiempo para clasificarlas, ni mucho menos para almacenarlas durante algún tiempo.

En la actualidad, el formato digital tiene sus variantes: un artículo debe poder leerse en una computadora de escritorio, una tableta, un teléfono celular y, en general, en cualquier tipo de pantalla que exista y que podrá existir.

Gracias al empeño y esfuerzo de quienes hicieron lo necesario, nuestra revista ginecología y obstetricia de México logró indizarse en la base de datos más codiciada: el Index Medicus (PubMed), enseguida en otras de igual o mayor relevancia (Scielo, Scimago, etc.). Las ediciones

más recientes dejaron de aparecer porque nos hace falta la versión con marcaje XML, tarea que hemos emprendido y que en breve la revista retornará a esa base de datos.

### ¿Qué es XML?

Es un acrónimo de *eXtensible Markup Language* = Lenguaje Extensible de Marcación. Es un lenguaje de licencia libre utilizado para almacenar datos en forma legible, que surgió por la necesidad de almacenar grandes cantidades de información. Se trata de un metalenguaje que permite definir reglas (o lenguajes, por eso es extensible) que especifican cómo marcar partes relevantes de un texto: palabras, frases, números, fórmulas, etc. En el texto de un artículo pueden marcarse sus elementos bibliográficos: título, autores, resumen, palabras clave, secciones, cuadros, figuras, citas, referencias, y más. El lenguaje XML no es sobre el formato, sino sobre el contenido. No es para usar letra tipo Arial de 12 puntos, centrado y justificado, sino para etiquetar que “este es el apellido del autor y aquí está el nombre”.

El marcaje XML se usa para estructurar precisamente todos los elementos de los textos en los procesos de edición contemporánea de los artículos y otros tipos de documentos. Cada elemento está definido por una etiqueta (tag).

Cuando se edita un artículo, los diseñadores gráficos, o las imprentas, transforman el archivo de Word en PDF o HTML, pero estos formatos no les dicen a las máquinas qué hay en su interior: quiénes lo escribieron, de donde son, ni de qué se trata.

Cuando un médico lee el título o el resumen de un artículo, rápidamente decide si le interesa, pero ¿cómo sabe una máquina cuál es el título del artículo que podría interesarle al lector? Pues, simplemente debe indicársele que busque artículos específicos, que contengan las mismas etiquetas. Expresado de otra manera se le dice a una base de datos o a un robot buscador de Google: “alerta, aquí, en esta parte del documento está el nombre de pila del autor del artículo y en este otro sus apellidos”.

El XML consiste en poner etiquetas para indicarles a las bases de datos en dónde se encuentran los contenidos: aquí está el resumen, aquí está la referencia bibliográfica, ésta es la institución donde trabaja el primer autor, el segundo autor se desempeña en esta otra y así sucesivamente. Toda esa información no es para los lectores, sino para las computadoras que buscan y muestran la información. Expresado de otra manera: los editores se esfuerzan para que los

lectores encuentren los artículos que buscan y los autores sean leídos y, sobre todo, citados por muchos otros que siguen una misma línea de investigación.

A partir de los textos marcados es posible extraer metadatos del artículo (título, autores, resumen, palabras clave, revista, volumen, número, paginación, si es original, de revisión o caso clínico) y construir su referencia bibliográfica. Esto significa que ya no es necesario teclear o copiar la referencia bibliográfica a partir de la primera página del artículo. El etiquetado XML asegura que la referencia bibliográfica es fiel, sin ningún tipo de error; por tanto, puede localizarse al momento.

Los editores de **ginecología y obstetricia de México** estamos trabajando no solo para reincorporarnos, en breve, a PubMed sino pertenecer a otras bases de datos, hasta conseguir que cualquier médico, esté en donde esté, encuentre fácilmente lo que se investiga en México, la epidemiología de sus enfermedades y los aportes de quienes han investigado nuevos métodos, nuevas enfermedades, nuevos tratamientos.

Enrique Nieto Ramírez