

Apuntes editoriales para autores

Metadatos

Publicado originalmente en Macpalxóchitl, octubre 2019

Dalila Fragoso Tejas

El uso de la red y la “obligatoriedad” de las tecnologías de la información aplicadas a la edición, distribución y difusión de los artículos científicos, requiere el uso de lenguaje de máquinas o algoritmos recopiladores de datos, boots, o cosechadores de datos, para hacer posible que la información publicada sea accesible a través de varios sistemas de información: directorios, catálogos e índices.

Seguro que usted se preguntará ¿por qué obligatoriedad? Las revistas académicas requieren estar incluidas en diferentes tipos de sistemas de información para lograr una mayor difusión de sus publicaciones. Estos sistemas amplifican la difusión y facilitan la búsqueda. Las bases de datos y catálogos ponen a disposición del público la información en extenso, así como datos relacionados con el estado de actualización, últimos números publicados, sus títulos y autores. Por otro lado, los catálogos e índices en general sólo muestran los datos generales de la publicación, los títulos de las publicaciones, pero no necesariamente incluyen resúmenes ni los trabajos en extenso, pero sí vinculan los artículos con la plataforma particular de cada revista. Estas plataformas recolectoras de información bibliográfica pueden ser universales (Web of Scineces o Ulrich) o regionales y especializadas para diferentes áreas del conocimiento (PERIODICA, REDALYC, LATINDEX), lo cual es muy útil, pues recopilan, almacenan y distribuyen información más precisa. Todo esto para facilitarnos las búsquedas de los materiales que más se apeguen a nuestros intereses. Por otra parte, incluyen información sobre las normas editoriales, datos generales, incluyendo los enlaces a su portal, donde los artículos pueden consultarse en acceso abierto. Es por eso que se vuelve obligatorio para los editores de las

revistas en línea, facilitar el acceso de la información de la revista a estos buscadores y cosechadores de información.

Los ejemplos más conocidos de bases de datos son: Scopus, Web of Sciences, Scielo, PubMed, Worldcat, REDALYC, entre otras, que incluyen acervos muy completos de contenidos tanto de publicaciones periódicas como libros de ciencias exactas y médicas. Ejemplo de catálogos: Dialnet plataforma de contenidos hispanos, especializado para publicaciones de ciencias sociales; Latindex, que ofrece un directorio de datos bibliográficos y de contacto con todas las revistas registradas, así como un catálogo compuesto por las revistas con más altos estándares de calidad, de acuerdo con sus criterios de inclusión en su portal; REDIB es una plataforma de agregación de contenidos científicos y académicos para Iberoamérica; BIBLAT de la UNAM, especializado en bibliografía de América Latina y el Caribe, se divide en CLASE (publicaciones en el área de sociales) y PERIODICA (especializada en el área de ciencias); DOAJ es un directorio que indexa publicaciones de alta calidad exclusivamente de acceso abierto, con revisión por pares; entre otras.

¿Por qué es necesario un lenguaje de máquinas? El lenguaje de máquinas está codificado en 0's y 1's. Para que dicho lenguaje sea legible para humanos, es necesario una interfase que transforme dichos códigos, un documento transformado a lenguaje humano puede ser un formato PDF. En ese sentido, un archivo en PDF no es un documento que las máquinas puedan leer y de donde puedan obtener información, ya que un PDF es un archivo "muerto". Es decir, un PDF es un texto sólo legible para humanos, ya no está codificado en 0's y 1's, es como una foto del lenguaje legible para humanos. Ha perdido su valor de rastreo para las máquinas. Por ello se hace necesario que todos los documentos en PDF sean transformados a un tipo de lenguaje que sea legible para los *boots* o cosechadores de información de los índices, catálogos y directorios, de tal forma que puedan llenar sus bases de datos con la información correcta.

¿Qué requieren esos cosechadores de información? Requieren que los datos en un documento estén ordenados específicamente siempre en el mismo orden y/o identificados con marcas especiales. La información que debe estar marcada es: nombre de la revista, ISSN, nombre del artículo, nombre de los autores, ORCID, número DOI, fecha aceptación de la publicación, fecha de publicación, editor asociado, texto del artículo en sí mismo,

cada una de las referencias bibliográficas, y en algunos casos las figuras y apéndices. Los archivos ordenados y marcados tienen un código universal o formato, que se llama XML, o HTML. Hay diferentes tipos de marcadores para documentos en XML. En el ámbito editorial, hay un patrón específico que las revistas pueden usar. Sin embargo, hay otras marcas XML para otros tipos de documentos, por ejemplo, las facturas de hacienda.

Todo lo que se sube a la red y es visible para todo el público, debe incluir información sobre su autor, una descripción, fecha de elaboración, permisos o tipo de derechos, características de la publicación, etc. Por eso es necesario que cuando usted sube un manuscrito al sitio de Botanical Sciences en OJS, incluya toda la información que le solicitan los formularios. En breve la revista Botanical Sciences, cambiará a una nueva versión del sistema, donde está contemplada la captura de información para los diferentes elementos que conforman los archivos complementarios del texto principal, índices, figuras, u otro tipo de archivos. En la nueva versión será necesario que se llenen todos los datos solicitados, para que una vez en la red, sean leídos los metadatos y la información tanto en nuestro sistema como en otros sistemas bibliográficos y plataformas de bases de datos, la información se preserve lo más completamente posible.