

Sustentabilidad de las iniciativas latinoamericanas de publicación de revistas científicas en acceso abierto utilizando el estándar XML JATS: el caso de SciELO

Eliana Guzmán-Useche,
Fernando Rodríguez-Contreras

Universidad de Los Andes – ULA, Venezuela

ARTICLE

Resumen

Objetivo: realizar el análisis de sustentabilidad de la iniciativa latinoamericana scielo.org, basado en dos aspectos primordiales, el económico y la necesidad de su existencia. Método: se efectúa la medición de las horas-hombre requeridas para publicar un artículo científico, usando su nueva metodología de publicación, denominada XML SPS. Se procesan 40 artículos publicados en revistas científicas de diferentes áreas del saber, pertenecientes a Brasil, Chile, Colombia, México y Venezuela, usando dicha metodología. Resultados: se obtiene que la preparación del documento tarda en promedio 35.0% del tiempo ($\sigma=9.72\%$) y la marcación tarda en promedio el 50.0% del tiempo ($\sigma=8.50\%$), siendo las secciones de referencias, agradecimientos y apéndices las que consumen más tiempo. Conclusiones: Esto indica que el modelo de publicación de SciELO es sustentable desde el punto de vista de la necesidad de su existencia, pero en el aspecto económico parece no ser completamente sustentable para todos los países de la región, porque es un proceso técnicamente complejo que no todos los editores son capaces de realizar, siendo necesaria la tercerización, lo cual exige un mayor presupuesto para las revistas y finalmente puede poner en riesgo la sustentabilidad y control de este modelo de publicación.

Palabras clave

Sustentabilidad ; Acceso abierto ; Latinoamérica ; Revista científica ; Redalyc ; SciELO ; JATS

Sustainability of Latin American initiatives that publish open access journals using the XML-JATS standard: the case SciELO

Abstract

Objectives: It is performed a sustainability analysis of the Latin American initiative scielo.org based on two main aspects, the economic and the need for its existence. Method: To achieve that, the required man-hours to publish a scientific paper were measured, using its new publication methodology called XML SPS. 40 articles published in scientific journals of different areas of knowledge are processed; journals from Brazil, Chile, Colombia, Mexico and Venezuela, using that methodology. Results showed that the preparation of the document takes on average 35% of the total time ($\sigma = 9.72\%$), while the tagging process takes on average 50% of the time ($\sigma = 8.50\%$), finally the sections corresponding to references, acknowledgments and appendices (back) takes the majority of the time. Conclusions: This analysis indicates that the SciELO publishing model is sustainable from the point of view of the need for its existence, but in the economic terms seems not to be fully sustainable for all countries in the region, because it is a technically complex process that not all publishers are able to perform, making outsourcing of that procedure to be necessary, which requires a greater budget for the journal that can jeopardize the sustainability and the control of this publishing model.

Keywords

Sustainability ; Open access ; Latin America ; Journal ; Redalyc ; SciELO ; JATS

1. Introducción

El presente trabajo es producto de un proyecto de investigación sobre “Acceso Abierto y Sustentabilidad en América Latina” que se lleva a cabo entre la Universidad de Los Andes, Venezuela, y la Universidad Autónoma del Estado de México, México. La permanencia en el tiempo de forma eficiente es el reto de los modelos de acceso abierto, lo que está determinado por la posibilidad de sustentabilidad de los dos principales actores que han adoptado dicho modelo en Latinoamérica: scielo.org (CNPq-Brasil) y redalyc.org (UAEM-México). Dicho proyecto analiza la sustentabilidad de los procesos emprendidos por cada uno de estos dos organismos, con la finalidad de poder argumentar, basados en evidencia, sobre las ventajas y desventajas de las estrategias emprendidas y proporcionar elementos a las comunidades de editores y académica para iniciar la discusión sobre el modelo de publicación latinoamericano.

El Journal Article Tag Suite (JATS) es un estándar de NISO Z39.96-2015, que define un conjunto de elementos y atributos XML para etiquetar el contenido completo de los artículos publicados en revistas científicas, con el fin de describir su estructura, semántica y metadatos. JATS es una continuación de la obra NLM de Archivo e Intercambio DTD iniciada en 2002 por el National Center for Biotechnology Information (NCBI, 2013). Dicho estándar ha sido adoptado por diferentes sistemas internacionales de publicación electrónica de revistas científicas, ya que es independiente de la plataforma tecnológica, convirtiéndolo en la norma por seguir, para así evitar problemas de interoperabilidad y trabajo extra en llevar a cabo conversiones de formato. Además permite que el archivo de cada artículo publicado en la revista vuelva a estar en manos del editor, usando el formato estándar que le permite su preservación digital y le garantiza su visibilidad a futuro. Esto es ya un paso más para lograr la sustentabilidad de la publicación de revistas científicas en acceso abierto.

En el presente trabajo se analiza la sustentabilidad del modelo SciELO desde la perspectiva técnica, es decir, del costo en horas-hombre que implica generar el contenido de una revista científica usando el XML SciELO Publishing Schema (XML SPS). Se decide comenzar por el análisis de SciELO, debido a que desde 2014 tiene disponible la herramienta tecnológica que permite generar los archivos usando el estándar JATS, mientras que Redalyc hizo el anuncio de comenzar a usar dicho estándar en la Reunión Nacional de Editores de Revistas Científicas Mexicanas, realizada en la Universidad de Chapingo del 27 al 29 de agosto de 2015. Además, mencionó que están desarrollando un programa informático dirigido a los editores de las revistas, que permitirá generar los contenidos usando dicho estándar de publicación electrónica. Una vez que la herramienta de Redalyc esté disponible y se puedan generar los archivos XML JATS, se realizará el correspondiente análisis de sustentabilidad.

Sería beneficioso para la región que Redalyc y SciELO definieran lineamientos comunes para crear y mejorar los procesos involucrados en la generación de los contenidos de las revistas científicas usando el estándar JATS, de manera que se logre la interoperabilidad en la región, se evite la duplicación de trabajo para los editores de las revistas, se minimicen los costos de la publicación en acceso abierto, de forma que se garantice la sustentabilidad de ambos proyectos. En principio, que estos dos metaeditores de la región adopten dicho estándar pareciera una buena noticia.

Es obvio que el acceso abierto tiene costo. Lo que plantea dicha filosofía es no trasladar el costo al usuario final, el cual es generalmente cubierto por fondos gubernamentales o del presupuesto de comunidades académicas. En este sentido, conocer los costos de la generación del archivo XML JATS permitirá avanzar en propuestas de sustentabilidad.

La sustentabilidad del acceso abierto ha sido analizada desde diferentes puntos de vista, como por ejemplo, la sustentabilidad de repositorios institucionales (Babini y Machin-Mastromatteo, 2015), de revistas académicas en acceso abierto (Public Knowledge Project [PKP], s/f), de la comunicación académica en América Latina (Alperin, Fischman y Willinsky, 2011), de los servicios (Crow, 2013) y de las bibliotecas universitarias (Babini, 2012), pero no ha sido analizada desde el punto de vista del aspecto técnico de la generación de los contenidos, el cual es importante para garantizar la interoperabilidad de la publicación en acceso abierto, específicamente de las revistas contenidas en SciELO y Redalyc.

La sustentabilidad de una iniciativa de publicación electrónica en acceso abierto primordialmente debe tomar en cuenta dos aspectos, el económico y la necesidad de que exista dicha iniciativa (Crow, 2013). Partiendo de esta premisa, no se discute que tanto Redalyc como SciELO son iniciativas de acceso abierto sustentables en cuanto al aspecto de la necesidad de su existencia para América Latina, debido a que son los dos líderes dominantes en la región (Alperin et. al., 2011).

Mucho se ha escrito sobre el modelo SciELO y el éxito obtenido en la región para hacer visible la producción científica latinoamericana, pero hasta el momento son pocos los trabajos existentes que permiten analizar el aspecto económico de la sustentabilidad de dicho modelo. Rozemblum, Bava, Unzurrunzaga y Banzato (2014) efectuaron un análisis de los costos y beneficios de incluir revistas universitarias en bases de datos de acceso abierto, haciendo una comparación entre Redalyc, SciELO y DOAJ, y obtuvieron como resultado que para un número de revista de 10 contribuciones, SciELO es 54.67 veces más caro que Redalyc y 82 veces más que DOAJ, por lo que concluyeron que del 100% del costo anual de incluir una revista en estos tres sistemas de acceso abierto, el 97.0% de los costos se los lleva SciELO, 2.0%, Redalyc y 1%, DOAJ. Esta gran diferencia de costos se debe a que SciELO hace el marcaje detallado de las referencias y ahora con su nueva metodología de marcación, XML SPS, agrega la marcación de los elementos del cuerpo del artículo. Por su parte, DOAJ y Redalyc únicamente generan los metadatos de la ficha bibliográfica sin hacer marcaje de referencias, aunque Redalyc realiza un proceso de normalización de instituciones para la generación de los indicadores de producción, coautoría y uso. De ahí la gran diferencia en los costos, debido a que son procesos distintos.

Por otra parte, Herrera, Ferracutti, Martínez y Mitzig (2011) presentan un gráfico sobre la distribución de los días de procesamiento de 14 números de diferentes revistas usando la metodología de SciELO, los cuales se dividen en días de generación del HTML (7 días en promedio, $\sigma = 2$ días) y días de marcación (7 días en promedio, $\sigma = 2.4$ días), y obtienen que se necesitan 14 días hábiles en promedio ($\sigma = 4$ días) para procesar un número de aproximadamente 12 contribuciones.

Los análisis actuales sobre la sustentabilidad de los modelos de publicación en acceso abierto apuntan a que se deben redirigir los recursos financieros que se invierten en el pago de suscripciones a revistas hacia el pago por servicios de publicación en acceso abierto (*Article Processing Charge*, APC). Dicho cambio en el modelo de negocio tiene mayor probabilidad de lograrse si se aborda como una iniciativa a gran escala de país o región (Velterop, 2015). No hay que dudar que ya se transita por el camino del cambio de modelo, al observar lo que están haciendo algunas editoriales, como IOP y la Royal Society of Chemistry, al adoptar un modelo de compensación que reduce el costo de suscripción a la revista, de acuerdo con la cantidad pagada en APC de los artículos publicados en ésta (Cop, 2015). Estos modelos son aplicables a las realidades de Estados Unidos y Europa, pero constituyen un riesgo en América Latina porque no existe la garantía de que los investigadores, instituciones o presupuestos de los proyectos de investigación puedan costear los APC (Babini y Machin-Mastromatteo, 2015).

Actualmente, los esfuerzos en Latinoamérica se deben enfocar en lograr la sustentabilidad de las dos grandes plataformas de publicación de revistas científicas en acceso abierto, Redalyc y SciELO; se debe comenzar por lograr la interoperabilidad entre éstas, sin perder de vista la inclusión de otras iniciativas de acceso abierto, como por ejemplo LA Referencia (Babini y Machin-Mastromatteo, 2015). Si bien es cierto que ambas plataformas de publicación han hecho desarrollos tecnológicos importantes y promueven el acceso abierto, aún falta por resolver algunos aspectos de interoperabilidad que permitan el intercambio de información (Swan, 2013). Esto debe resolverse lo antes posible, ya que es uno de los aspectos que garantizan la sustentabilidad de ambas plataformas y contribuyen a cerrar la brecha que nos separa de las iniciativas anglosajonas y europeas.

El presente análisis de sustentabilidad hace uso de un método experimental – replicable – que permite medir el costo en horas-hombre que se emplean en publicar electrónicamente un artículo de una revista científica, usando la actual metodología del modelo SciELO, lo que posibilita evaluar cómo este metaeditor está aplicando el estándar JATS, la eficiencia de su herramienta de marcación y las características del archivo XML resultante.

2. Scientific Electronic Library Online: SciELO

En 1998 comienza el funcionamiento regular del Programa Scientific Electronic Library Online (SciELO) en Brasil, el cual se había iniciado un año antes como un proyecto piloto que tenía como objetivo desarrollar la infraestructura necesaria para indizar y publicar electrónicamente en Internet revistas científicas y, de esta manera, aumentar su visibilidad, uso e impacto y, por ende, el de la investigación científica en ésta contenida (Packer et. al., 2014a).

Según Packer et. al. (2014a), SciELO tiene dos funciones principales: a) la función de indización, que consiste en utilizar criterios de calidad editorial de las revistas con el fin de ampliar su cobertura y emplear un sistema en línea que le permita medir el rendimiento en función de la cantidad de descargas y el uso de indicadores bibliométricos basados en citas; y b) la función de publicación en la web que le permite a las revistas científicas operar bajo la filosofía del acceso abierto, obteniendo los beneficios típicos de la navegación y

búsqueda de información, además de garantizar la interoperabilidad de los contenidos con la finalidad de hacer una distribución masiva de los metadatos en la web.

En Latinoamérica, 11 países han adoptado el modelo SciELO como parte de sus políticas públicas, para dar visibilidad y publicar los contenidos de sus revistas científicas; también lo han hecho España, Portugal y Sudáfrica. En su totalidad, integran lo que se conoce como la Red SciELO, la cual cuenta con 573,525 artículos científicos (SciELO, 2015d). Pero cada país es el responsable de la coordinación, financiación y funcionamiento de su colección nacional, razón por la cual debe ser importante que los encargados de tomar la decisión sobre la adopción del modelo SciELO conozcan los requerimientos técnicos y económicos necesarios que dicha decisión implica, en otras palabras, que evalúen la sustentabilidad de este modelo.

Desde sus inicios, SciELO utiliza los textos de los artículos científicos en formato HTML (Hyper Text Markup Language), con la finalidad de preservar la integridad del texto original, y se marca según la estructura SGML SciELO. Desde 2013, SciELO decide hacer uso de un nuevo formato para almacenar el texto completo de los artículos, el XML: eXtensible Markup Language (Santos y Packer, 2014).

En 2014, SciELO anunció el lanzamiento del nuevo esquema de publicación de sus contenidos, el XML SPS, el cual se basa en adoptar las especificaciones de NISO JATS Journal Publishing DTD versión 1.0 (NCBI, 2013), además del Estilo SciELO, para generar el archivo XML que debe enviar el editor de la revista al administrador del sitio SciELO del país correspondiente, para su publicación electrónica (SciELO, 2015a). SciELO agrega al estándar el uso de algunas etiquetas JATS para identificar de forma más detallada las afiliaciones de los autores, datos de financiamiento de la investigación y la marcación detallada de las referencias bibliográficas con la finalidad de generar sus indicadores bibliométricos (Packer et. al., 2014b).

Se puede observar en el sitio SciELO de Brasil (<http://www.scielo.br>) que para publicar los números de las revistas correspondientes a 2015 es obligatorio el uso del esquema XML SPS (Packer et. al., 2014b). En el resto de los sitios SciELO de la red no existe tal exigencia y por lo tanto se adecuan a él en la medida de sus posibilidades. Por ejemplo, el sitio SciELO-México está usando este esquema y ha comenzado la publicación del XML JATS desde finales de 2015; al igual ha avanzado de forma importante Costa Rica. Los otros países irán sumándose con mayor o menor rapidez al proceso, de acuerdo con sus posibilidades técnicas y financieras.

Al analizar los tres componentes del modelo SciELO (SciELO, s/f) se observa que la implementación de un sitio SciELO recae fundamentalmente en una institución pública nacional; es decir, debe contar con el apoyo gubernamental para garantizar la puesta en marcha, el funcionamiento y la sustentabilidad en el tiempo de la publicación electrónica de las revistas científicas, ya que se requiere del trabajo de un recurso humano capacitado y el empleo de una plataforma tecnológica adecuada para el procesamiento y almacenamiento de datos, mantenimiento y publicación de los resultados.

3. Estructura jerárquica del XML JATS

Cuando se genera un archivo en formato XML JATS está conformado por etiquetas XML contenidas en el estándar JATS. Estos archivos tienen la extensión .xml y su contenido es un árbol de jerarquía de etiquetas, las cuales encierran los datos bibliométricos del artículo científico. Por lo general, los identificadores de las etiquetas están relacionados con el dato que se va a marcar o identificar, por ejemplo, la etiqueta <article-title> se usa para marcar el título del artículo. A su vez, las etiquetas pueden tener atributos, que sirven para enriquecer el contenido de la etiqueta; siguiendo con el ejemplo, esta misma etiqueta puede usar el atributo xml:lang con el valor "es" para indicar que el idioma del artículo es español <article-title xml:lang="es">.

Además, las etiquetas en el XML tienen jerarquía, es decir, ciertas etiquetas pueden estar contenidas dentro de otras etiquetas, siempre y cuando la definición del estándar lo permita. De esta manera, los archivos XML típicamente forman un árbol de jerarquías.

Para el caso de los artículos científicos, la etiqueta principal es <article>, que a su vez se divide en cinco componentes (NCBI, 2012), siendo los tres primeros los que se usan con mayor frecuencia:

1. <front>: contiene los metadatos, tanto de la revista como del artículo.
2. <body>: contenido textual y gráfico del artículo.

3. <back>: lista de referencias, agradecimientos, apéndices y glosario.
4. <floats-group>: elementos flotantes que utilizan algunos editores, tales como figuras y tablas, que son referenciadas en el texto o en las referencias.
5. <response> o <sub-article>: <reponse> se utiliza para identificar un comentario del propio artículo. <sub-article> se utiliza para identificar la traducción del artículo en algún otro idioma.

Para generar el archivo XML JATS de un artículo científico se debe marcar la información contenida en el front, que es la utilizada por los recolectores de metadatos; la del body para ver el texto, imágenes, ecuaciones y/o tablas a través del visor; y en el back para generar los indicadores de citación.

4. Materiales y método

4.1. Selección de la muestra de revistas

Para medir la cantidad de horas-hombre requeridas para publicar electrónicamente un artículo de una revista científica, usando la actual metodología de SciELO, se decide clasificar los artículos científicos que son parte del experimento, según su tipo (editorial o presentación, artículo original, reseñas, ensayos, avances de investigación, entre otros) y área de conocimiento, debido a la diferencia en la cantidad de texto, ecuaciones matemáticas, imágenes, cuadros, figuras, tablas, notas y referencias bibliográficas que existen entre los artículos científicos de las diferentes áreas. Algunos de éstos pueden presentar una mayor cantidad de texto pero menor cantidad de tablas, como por ejemplo en los artículos científicos contenidos en las revistas del área de las ciencias sociales, comparadas con los de las áreas de ciencias exactas o naturales.

Con el objetivo de analizar una muestra de revistas que represente la realidad latinoamericana, se seleccionan 40 artículos publicados en revistas científicas latinoamericanas, haciendo uso del ranking de revistas del Scimago Journal & Country Rank de 2014 (SCImago, 2007), y bajo el criterio de seleccionar artículos publicados en revistas que posean valores altos, medios y bajos del Índice de Impacto (SJR). En el cuadro 1 se muestran las revistas seleccionadas por área y país y los datos específicos del artículo por procesar. El método de selección consiste en escoger la revista, por área de conocimiento, que tenga el mayor SJR de su país de edición.

Además, en esta selección de artículos científicos se toma en cuenta el área de conocimiento de la revista y que sea de publicación reciente. En cuanto a la clasificación de las revistas científicas según su área de conocimiento, se usa la de SciELO, para ser consecuentes con su forma de trabajo y poder usar las cifras que publican como parte de los Indicadores Bibliométricos de la red SciELO (SciELO, 2014). En dicha clasificación se destaca que cuatro áreas pertenecen a las ciencias sociales y humanas, mientras que las otras cuatro son de las ciencias naturales y exactas. La clasificación por usar es la siguiente:

- a) Ciencias agrícolas.
- b) Ciencias biológicas.
- c) Ciencias de la salud.
- d) Ciencias exactas y de la tierra.
- e) Ciencias humanas.
- f) Ciencias sociales aplicadas.
- g) Ingenierías.

h) Lingüística, letras y artes.

4.2. Materiales

Para procesar los artículos se usa la aplicación SciELO PC Programs 4.0.090-trial, disponible en el sitio de descargas de productos SciELO (<ftp.scielo.br>). Esta versión es la que implementa el uso de los estándares de JATS para generar los archivos XML.

Se realiza la configuración de SciELO PC Programs según los requerimientos establecidos (SciELO, 2015b), haciendo uso de JAVA versión 1.8.0_60, PYTHON versión 2.7.10 y PILLOW versión 1.1.7. Cabe señalar que dentro del conjunto de aplicaciones incluidas en SciELO PC Programs se encuentra el Markup, que es una macro para Microsoft Word® y se usa para el marcaje de los artículos; por lo tanto, los documentos por marcar deben estar en dicho formato (.doc o docx). En este trabajo se utiliza la versión de Microsoft Word® 2010.

Cuadro 1 - Revistas seleccionadas para la marcación usando la metodología de SciELO

Área	País	Brasil	Chile	Colombia	México	Venezuela
Ciencias agrícolas		<i>Revista Ciência Agronômica</i> 2015, 46(4), 669-678 pp., 29482 caracteres	<i>Chilean Journal of Agric. Research</i> 2015, 75(3), 320-327 pp., 30544 caracteres	<i>Rev. Colombiana Ciencias Pecuarias</i> 2015, v28n3, 238-246 pp., 17750 caracteres	<i>Rev. Mexicana de Biodiversidad</i> 2015, 86(1), 59-71 pp., 41247 caracteres	<i>Bioagro</i> 2014, 26(1), 39-48 pp., 24382 caracteres
	SJR	0.685	0.347	0.330	0.252	0.199
Ciencias biológicas		<i>Neotropical Ichthyology</i> 2015, 13(3), 547-556 pp., 39946 caracteres	<i>Biological Research</i> 2015, 48, 1-8 pp., 28300 caracteres	<i>Caldasia</i> 2015, 37(1), 107-123 pp., 38197 caracteres	<i>Ciencias Marinas</i> 2015, 41(1), 21-32 pp., 51100 caracteres	<i>Acta Botánica Venezolana</i> 2013, 36(1), 15-38 pp., 40524 caracteres
	SJR	0.542	0.474	0.246	0.486	0.105
Ciencias de la salud		<i>Revista Saúde Pública</i> 2014, 48(6), 925-930 pp., 16728 caracteres	<i>Terapia Psicológica</i> 2015, 33(2), 81-92 pp., 40797 caracteres	<i>Biomédica: Rev. Inst. Nac. de Salud</i> 2015, 35(2), 204-211 pp., 25549 caracteres	<i>Annals of Hepatology</i> 2015, 14(6), 881-887 pp., 19494 caracteres	<i>Archivos Lat. de Nutrición</i> 2015, 65(3), 152-157 pp., 18234 caracteres
	SJR	0.696	0.343	0.252	0.654	0.243
Ciencias exactas y de la tierra		<i>Revista Brasileira de Ciência do Solo</i> 2015, 39(4), 1068-1078 pp., 40367 caracteres	<i>Andean Geology</i> 2015, 42(3), 329-348 pp., 55615 caracteres	<i>Rev. Colombiana de Química</i> 2014, 43(3), 18-24 pp., 22310 caracteres	<i>R. M. Astronomía y Astrofísica</i> 2015, 51(1), 101-118 pp., 38765 caracteres	<i>Rev. de la Facultad de Agronomía</i> 2015, 31(3), 325-340 pp., 37810 caracteres
	SJR	0.778	0.772	0.132	1.294	0.128
Ciencias humanas		<i>Psicologia e Sociedade</i> 2015, 27(3), 548-557 pp., 34666 caracteres	<i>Chungara</i> 2015, 47(3), 1-12 pp., 28047 caracteres	<i>Universitas Psychologica</i> 2015, 14(1), 15-25 pp., 27130 caracteres	<i>R. M. de Análisis de la Conducta</i> 2015, 41(2), 19-43 pp., 53206 caracteres	<i>Utopía y Praxis Latinoamericana</i> 2015, 20(68), 79-102 pp., 62624 caracteres
	SJR	0.362	0.342	0.264	0.211	0.149
Ciencias sociales aplicadas		<i>Cadernos PAGU</i> 2015, (44), 423-452 pp., 49033 caracteres	<i>Political Criminal</i> 2015, 10(19), 159-191 pp., 80940 caracteres	<i>Rev. Colombiana de Estadística</i> 2015, 38(1), 107-121 pp., 21039 caracteres	<i>RELIME</i> 2015, 18(1), 49-76 pp., 54487 caracteres	<i>Espacios</i> 2015, 36(20), 1-14 pp., 29233 caracteres
	SJR	0.384	0.391	0.300	0.319	0.115

SJR: SCImago Journal Rank.

Fuente: elaboración propia.

Área	País	Brasil	Chile	Colombia	México	Venezuela
Ingenierías		<i>Brasilian Journal of Chemical Engineer.</i> 2015, 32(3), 683-698 pp., 49153 caracteres 0.389	<i>Maderas: Ciencia y Tecnología</i> 2015, 17(3), 517-524 pp., 16871 caracteres 0.357	<i>DYNA</i> 2015, 82(192), 78-84 pp., 19212 caracteres 0.228	<i>J. of Applied Res. and Technology</i> 2015, 13(3), 409-415 pp., 14962 caracteres 0.279	<i>Rev. Lat. Metalurgia y Materiales</i> 2015, 35(2), 201-212 pp., 27641 caracteres 0.148
Lingüística, letras y artes		<i>Calidoscopio</i> 2015, 13(2), 189-200 pp., 53796 caracteres 0.182	<i>RLA</i> 2015, 53(1), 131-147 pp., 25455 caracteres 0.212	<i>Palabra Clave</i> 2015, 18(2), 341-373 pp., 46218 caracteres 0.172	<i>Estudios de Cultura Maya</i> 2015, 46, 103-118 pp., 30226 caracteres 0.102	<i>Boletín de Lingüística</i> 2013, XXV(39-40), 187-216 pp., 46874 caracteres 0.102

SJR: SCImago Journal Rank.

Fuente: elaboración propia.

Para el procesamiento de los artículos seleccionados se siguen las directrices de SciELO Publishing Schema (SPS) versión 1.3, publicados en la Guía de Uso de los Elementos y Atributos para Documentos en XML (SciELO, 2015a), en conjunto con la consulta del grupo de Google SciELO XML - SciELO Publishing Schema (SciELO, 2013).

Otro aspecto importante en este experimento es la experiencia del recurso humano que aplica la metodología de SciELO para la marcación de los artículos científicos seleccionados. En el presente análisis las personas que realizan el proceso de marcaje cuentan con 10 años de experiencia en la marcación de artículos usando la metodología SciELO, siendo capaces de:

- Clasificar un artículo científico según su tipo, es decir, si es un editorial o presentación, artículo original, reseñas, ensayos, avances de investigación, entre otros.
- Identificar los elementos de un artículo científico, como son los títulos, autores, resúmenes y palabras clave, que representan el encabezado; el cuerpo del artículo; las notas y las referencias bibliográficas.
- Reconocer los diferentes tipos de referencias, ya sean citas de libros, capítulos de libros, artículos de revistas, páginas web, documentos de un congreso, tesis, artículos de periódicos, videos, entre otros.

Además, las personas que ejecutan el trabajo cuentan con conocimientos en el tratamiento de un texto para realizar la preparación inicial del artículo científico, tratamiento de las imágenes, HTML, XML y JATS.

En el experimento es fundamental la utilización de personal con amplia experiencia en el uso del modelo de marcación SciELO. Por un lado, la herramienta no es sencilla y no ha sido prácticamente modificada desde su inicio, lo que exige –para que el proceso sea lo más eficiente y exitoso posible– que sea manejado por personal con una alta capacitación. De hecho, esta situación podría ser una crítica al modelo y un problema desde la perspectiva de la sustentabilidad, pero en este caso se hace énfasis en la cuestión tiempo total de procesamiento.

5. Resultados

Al aplicar la metodología descrita se obtienen los tiempos de procesamiento de un artículo científico, usando la metodología de SciELO, XML SPS. Dicha metodología se implementa en los siguientes pasos:

- **Paso 1.** Creación de la revista usando el Title Manager: la metodología de SciELO exige que para procesar cualquier revista, la primera vez se debe ingresar la información básica de la revista: ISSN, misión, área, descriptores, idioma del texto y resumen, periodicidad, año inicial, volumen inicial, número inicial, institución editora de la revista, dirección, correo electrónico, tipo de licencia Creative Commons, entre otra información. Este paso se hace una sola vez para todos los números.
- **Paso 2.** Creación de las secciones (Title Manager): se deben crear las secciones que usa cada número de la revista para generar la tabla de contenidos. Si el número por procesar tiene una sección que no ha sido registrada antes, se debe proceder a hacerlo.
- **Paso 3.** Creación del número de la revista por ser procesado (Title Manager): se debe crear el número de la revista que va a ser publicado, indicando el volumen, número, secuencia del número en el año de publicación, cantidad de artículos, temporada, entre otros datos.
- **Paso 4.** Preparación del documento: esta preparación comprende dos procesos; en el primero, se debe generar el artículo en un archivo de Microsoft Word® (.doc o docx), ya que así lo exige el proceso de marcación usando el Markup de SciELO. En este trabajo se generan dichos archivos a partir de un archivo PDF, aunque bien podría realizarse tomando como base un documento en HTML (siempre y cuando se tenga el documento en PDF almacenado en la carpeta correspondiente). Este procesamiento comprende el tratamiento del texto, notas al pie, imágenes, tablas, cuadros, figuras y ecuaciones que tenga el artículo. El segundo proceso consiste en arreglar el documento según los requerimientos de SciELO, para comenzar la marcación, como lo son el orden del DOI, sección a la que pertenece el documento, títulos del artículo, autores, afiliaciones, tamaño de la letra para los títulos de las secciones y subsecciones, interlineado tanto en el cuerpo del artículo como en las referencias e inclusión de las imágenes (SciELO, 2015c).
- **Paso 5.** Marcación usando el Markup: consiste en la marcación del texto del documento usando las etiquetas de SciELO. En este caso se usa la metodología descrita por SciELO de su SPS versión 1.3 (SciELO, 2015a), con la finalidad de generar estos documentos en formato XML JATS, según el esquema de SciELO.
- **Paso 6.** Generación del XML: una vez que el documento ha sido completamente marcado, se debe generar el archivo XML, el cual se acompaña con los archivos de imágenes y el documento en formato PDF.
- **Paso 7.** Depuración de errores del XML: al generar el archivo XML es posible que existan errores en la marcación del documento, lo que requiere revisar los informes de errores que se generan y depurarlos.

5.1. Medición de tiempos

En la figura 1 se presentan los tiempos de procesamiento utilizando la metodología SciELO para generar los archivos XML JATS. Al desglosar por pasos los tiempos de procesamiento de los artículos presentados en la figura 1 se observa que:

- Crear la revista, lo cual se hace una sola vez, toma entre 4 y 10 minutos.

- Crear las secciones usadas en el número que se está procesando toma entre 1 y 2 minutos.
- Crear el número a ser marcado toma entre 1 y 2 minutos.
- La preparación del documento para ser marcado toma entre 20 y 95 minutos.
- La marcación es el proceso que más tiempo ocupa, entre 20 y 110 minutos.
- Generar el archivo XML al finalizar el marcado toma entre 1 y 6 minutos (si no hay errores en la marcación es más rápido, pero si hay errores de marcación puede tardar hasta 10 minutos).
- Depurar los errores toma entre 0 (si no hay ningún error) y 15 minutos (siempre y cuando quien procesa tenga conocimientos de las etiquetas de SciELO y de XML).

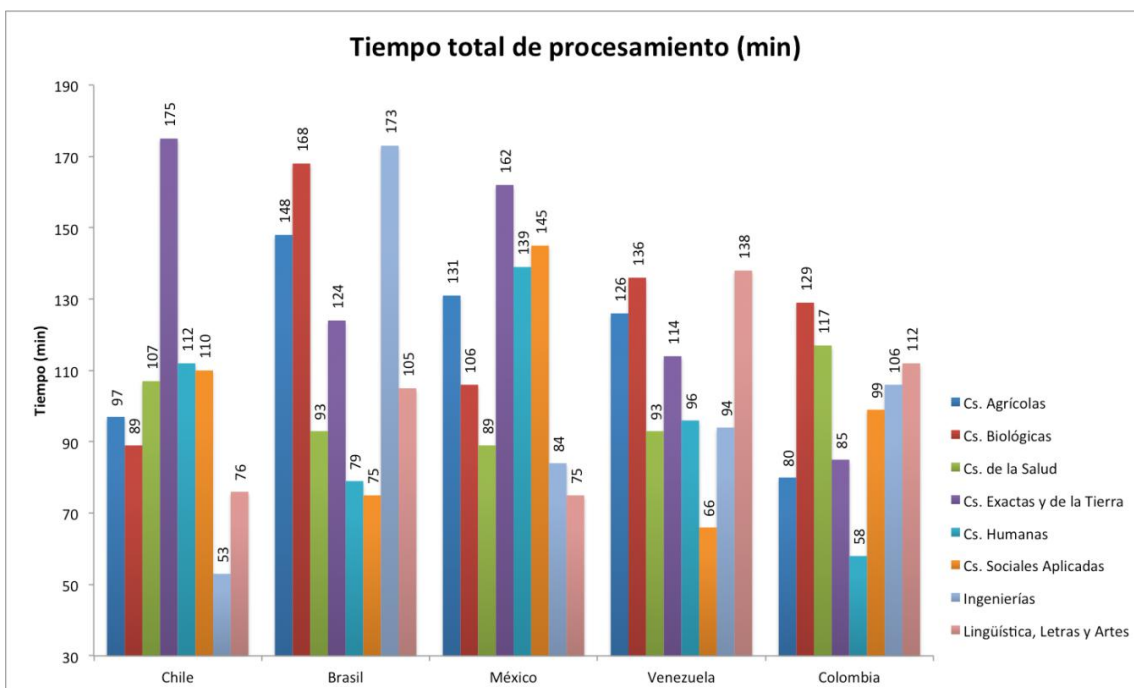


Figura 1 - Tiempos totales de procesamiento usando la metodología de SciELO

En el cuadro 2 se presenta la razón de la cantidad total de caracteres del documento entre el tiempo empleado en su procesamiento. Se observa que la cantidad de caracteres que tenga el artículo no afecta el tiempo total de procesamiento, por ejemplo, si se compara la razón en el área de Ciencias de la Salud, se observa que para Chile (381 a 1) dicha razón es un poco más del doble que para Brasil (180 a 1) y el tiempo total de procesamiento de Chile (107 min.) no es más del doble que el de Brasil (93 min.).

Cuadro 2 - Razón caracteres/tiempo usando la metodología de marcación de SciELO

Área	Brasil	Chile	Colombia	México	Venezuela
Ciencias Agrícolas	199:1	315:1	222:1	315:1	194:1
Ciencias Biológicas	238:1	318:1	296:1	482:1	298:1
Ciencias de la Salud	180:1	381:1	218:1	219:1	196:1
Ciencias Exactas y de la Tierra	326:1	318:1	262:1	239:1	332:1
Ciencias Humanas	439:1	250:1	468:1	383:1	652:1
Ciencias Sociales Aplicadas	654:1	736:1	213:1	376:1	443:1
Ingenierías	284:1	318:1	181:1	178:1	294:1
Lingüística, Letras y Artes	512:1	335:1	413:1	403:1	340:1

Fuente: elaboración propia.

6. Análisis

6.1. Herramienta de Marcación de SciELO: Markup

Como se mencionó, para llevar a cabo el Paso 5, la marcación de un artículo científico usando la metodología de SciELO, se debe usar la aplicación de escritorio denominada Markup, la cual es una macro desarrollada para Microsoft Office Word®, que permite identificar los elementos bibliográficos en el texto de acuerdo con los lineamientos establecidos en el DTD de SciELO. En los inicios de SciELO, el DTD estaba definido para documentos de tipo article (artículos originales) y text (editorial, presentación o reseña). Actualmente se usa un nuevo DTD de SciELO denominado doc, el cual contiene los elementos necesarios para generar el XML según el SPS (SciELO, 2015e).

La herramienta Markup está compuesta por tres barras de menú, la primera es una barra de comandos (guardar, generar el XML, visualizar los informes de errores, entre otros), la segunda contiene 25 botones que corresponden a las etiquetas que permiten identificar diferentes elementos del body, mientras que la tercera contiene 20 botones para la marcación, tanto de los elementos del front y back. Como el XML JATS se estructura de una forma jerárquica, la mayoría de estos botones permiten observar todas las etiquetas que pueden ser incluidas dentro de la etiqueta primaria; por ejemplo, el botón refs que se usa para marcar la lista de referencias posee 5 niveles de menús, y en este caso el tercer nivel posee 32 botones, correspondientes a las etiquetas necesarias para marcar cada referencia. Es importante indicar que los nombres de los botones en la mayoría de los casos no son los mismos identificadores de las etiquetas JATS.

El uso de esta herramienta de marcación requiere un conocimiento especializado, no sólo en la identificación de los datos bibliográficos de un artículo científico, sino del XML, ya que una vez finalizada la marcación se debe generar el archivo XML JATS y cuando se presentan errores es necesario conocer las etiquetas y su respectivo uso para poder corregirlos. Además, se requiere un editor de XML, porque el Markup no está diseñado para tal tarea, y al tratar de corregir los errores no es muy útil.

Al analizar el Markup, se observa que es una aplicación que en sus casi 20 años de creación ha sido modificada de acuerdo con las necesidades de SciELO y con las diferentes versiones de Microsoft Office Word®, pero que realmente no ha tenido un cambio significativo en su diseño, que le permita ser más amigable con el usuario final y hacer uso de las nuevas herramientas tecnológicas disponibles. Vale la pena preguntarse si una mejora sustancial de esta herramienta de marcación reduciría los tiempos de procesamiento y por lo tanto los costos actuales. Hay que esperar a que Redalyc libere la herramienta de marcación y generación del XML JATS que está desarrollando, para verificar si es posible una reducción significativa de estos tiempos.

6.2. Comparación de tiempos entre revistas de Brasil, Chile, Colombia, México y Venezuela, pertenecientes a una misma área del conocimiento

Para poder comparar los tiempos medidos en cada uno de los seis pasos descritos anteriormente, se calculan sus correspondientes porcentajes y, como se observa en el cuadro 3, los tiempos correspondientes al paso 4 (preparación del documento) y al paso 5 (marcación usando el XML SPS) son los que consumen la mayor parte del tiempo de procesamiento de un artículo científico sin importar el área del conocimiento al que pertenezca. El paso 4 tarda en promedio 35.0% ($\sigma=9.72\%$) del tiempo, mientras que el paso 5 tarda en promedio el 50.0% ($\sigma=8.50\%$) del tiempo. Ambos pasos en conjunto ocupan entre el 70.0% y 90.0% del tiempo total de procesamiento. Estos porcentajes son muy similares a los obtenidos por Rozemblum et. al. (2014: 11).

Cuadro 3 - Porcentajes de tiempo de procesamiento en la preparación del documento (paso 4), en la marcación (paso 5) y su suma

Área	Brasil		Chile		Colombia		México		Venezuela	
Ciencias Agrícolas	20.3	77.7	27.8	85.5	28.8	80.1	29.0	87.8	32.5	81.7
	57.4		57.7		51.3		58.8		49.2	
Ciencias Biológicas	23.8	78.6	36.0	86.6	24.0	89.1	32.1	85.9	26.5	76.5
	54.8		50.6		65.1		53.8		50.0	
Ciencias de la Salud	45.7	83.8	36.4	88.7	35.0	87.1	28.1	91.0	43.0	86.0
	38.1		52.3		52.1		62.9		43.0	
Ciencias Exactas y de la Tierra	28.2	82.2	24.0	85.1	30.6	82.4	48.8	93.9	44.7	82.4
	54.0		61.1		51.8		45.1		37.7	
Ciencias Humanas	39.2	79.7	46.4	89.3	41.4	88.0	34.5	89.0	62.5	80.2
	40.5		42.9		46.6		54.5		17.7	
Ciencias Sociales Aplicadas	38.7	81.4	40.9	85.4	50.5	86.9	31.0	86.2	34.8	80.3
	42.7		44.5		36.4		55.2		45.5	
Ingenierías	38.2	87.9	28.3	77.4	45.3	91.5	40.5	81.0	44.7	88.3
	49.7		49.1		46.2		40.5		43.6	
Lingüística, Letras y Artes	43.3	84.6	44.7	78.9	34.8	90.2	54.7	89.4	58.0	89.2
	41.3		34.2		55.4		34.7		31.2	
Promedio	34.7	81.99	35.58	84.63	36.30	86.90	37.32	87.99	43.34	83.08
	47.3		49.06		50.60		50.67		39.73	
Desviación estándar	9.36	3.39	8.26	4.31	8.88	3.90	9.79	3.84	12.30	4.38
	7.53		8.56		8.24		9.65		10.80	

Fuente: elaboración propia.

Leyenda:

% Paso 4	% Paso 4 + % Paso 5
% Paso 5	

6.3. Estimación del tiempo promedio de procesamiento de un número completo

En el cuadro 4 se presentan los tiempos promedios de procesamiento de un artículo científico, usando la metodología de SciELO para generar su correspondiente archivo de XML JATS, según su área del conocimiento.

Cuadro 4 - Tiempos promedio de procesamiento de un artículo científico

Área	Mediana (min)	Promedio (min)	Desviación estándar (min)
Ciencias Agrícolas	126	116.4	27.41
Ciencias Biológicas	129	125.6	30.17
Ciencias de la Salud	93	99.8	11.80
Ciencias Exactas y de la Tierra	124	132.0	36.56
Ciencias Humanas	96	96.8	30.97
Ciencias Sociales Aplicadas	99	99.0	31.23
Ingenierías	94	102	44.29
Lingüística, Letras y Artes	105	101.2	26.90

Fuente: elaboración propia.

Con estos tiempos medidos se puede calcular la cantidad estimada de tiempo que toma procesar un número completo de una revista, lo cual dependerá de la cantidad de artículos que se publiquen en dicho número. Por ejemplo, si es un número de una revista del área de Ciencias Sociales Aplicadas que publica 12 artículos, su tiempo total de procesamiento usando la metodología de SciELO es aproximadamente de 20 horas; pero si la revista es del área de Ciencias Exactas y de la Tierra, y tiene la misma cantidad de artículos, el tiempo será de 27 horas. Es importante observar la dispersión que tiene el tiempo promedio de procesamiento de un artículo cuando se pretende hacer dichas estimaciones.

En el presente análisis, los tiempos de procesamiento del paso 4 (preparación del documento) y del paso 5 (marcación) son mucho menores que los obtenidos por Herrera et. al (2011), quien indica que requiere de 14 días en promedio ($\sigma = 4$ días) para procesar un número de 12 contribuciones con la metodología de SciELO, mientras que los tiempos máximos obtenidos en el presente estudio no llegan a ser mayores de 13 horas (menos de dos días de trabajo de 8 horas cada uno).

6.4. Contenido del artículo

Un aspecto importante a tomar en cuenta, ya que cambia en gran medida el tiempo total requerido en la marcación de un artículo científico, es la cantidad de referencias, figuras, tablas o cuadros, ecuaciones, páginas y notas al pie de página que dicho artículo tenga.

Como se observa en el cuadro 5, se puede determinar que la diferencia en el tiempo de procesamiento se debe en gran medida a la cantidad de referencias bibliográficas que tenga un artículo científico. Por ejemplo, al comparar los artículos del área de Ciencias Biológicas se observa que a pesar de que el artículo venezolano tiene muchas más páginas (24 p.) que el brasileño (10 p.), la diferencia en el tiempo de procesamiento es de 32 minutos (figura 1), debido a que este último tiene 70 referencias, mientras que el primero tiene 39. Esta misma relación entre el tiempo total de procesamiento y la cantidad de referencias es reseñada por Herrera et. al 2011, cuando establece que el tiempo de marcado depende de la cantidad y calidad de las referencias del artículo.

Cuadro 5 - Desglose de las características de los artículos científicos procesados usando la metodología de SciELO

Área	Autores	Idiomas	Páginas	Referencias	Notas	Figuras	Tablas
Ciencias Agrícolas	Brasil: 3	Brasil: 2	Brasil: 10	Brasil: 41	Brasil: 0	Brasil: 4	Brasil: 2
	Chile: 4	Chile: 1	Chile: 8	Chile: 42	Chile: 1	Chile: 3	Chile: 2
	Colombia: 6	Colombia: 3	Colombia: 9	Colombia: 21	Colombia: 0	Colombia: 0	Colombia: 8
	México: 4	México: 2	México: 13	México: 47	México: 2	México: 4	México: 3
	Venezuela: 1	Venezuela: 2	Venezuela: 10	Venezuela: 17	Venezuela: 0	Venezuela: 5	Venezuela: 8
Ciencias Biológicas	Brasil: 4	Brasil: 2	Brasil: 10	Brasil: 70	Brasil: 0	Brasil: 2	Brasil: 3
	Chile: 5	Chile: 1	Chile: 8	Chile: 27	Chile: 3	Chile: 0	Chile: 5
	Colombia: 2	Colombia: 2	Colombia: 17	Colombia: 36	Colombia: 0	Colombia: 3	Colombia: 6
	México: 3	México: 2	México: 12	México: 29	México: 0	México: 4	México: 0
	Venezuela: 2	Venezuela: 2	Venezuela: 24	Venezuela: 39	Venezuela: 0	Venezuela: 6	Venezuela: 2
Ciencias de la Salud	Brasil: 7	Brasil: 2	Brasil: 6	Brasil: 23	Brasil: 1	Brasil: 3	Brasil: 3
	Chile: 6	Chile: 2	Chile: 12	Chile: 29	Chile: 2	Chile: 1	Chile: 2
	Colombia: 3	Colombia: 2	Colombia: 8	Colombia: 21	Colombia: 0	Colombia: 1	Colombia: 2
	México: 5	México: 1	México: 7	México: 23	México: 1	México: 2	México: 4
	Venezuela: 6	Venezuela: 2	Venezuela: 6	Venezuela: 23	Venezuela: 0	Venezuela: 2	Venezuela: 2
Ciencias Exactas y de la Tierra	Brasil: 7	Brasil: 2	Brasil: 11	Brasil: 61	Brasil: 1	Brasil: 4	Brasil: 1
	Chile: 2	Chile: 2	Chile: 20	Chile: 95	Chile: 0	Chile: 9	Chile: 2
	Colombia: 9	Colombia: 3	Colombia: 7	Colombia: 18	Colombia: 0	Colombia: 5	Colombia: 2
	México: 3	México: 2	México: 18	México: 53	México: 7	México: 7	México: 7
	Venezuela: 5	Venezuela: 2	Venezuela: 15	Venezuela: 24	Venezuela: 1	Venezuela: 0	Venezuela: 4
Ciencias Humanas	Brasil: 2	Brasil: 3	Brasil: 10	Brasil: 23	Brasil: 2	Brasil: 0	Brasil: 4
	Chile: 2	Chile: 2	Chile: 12	Chile: 27	Chile: 1	Chile: 4	Chile: 3
	Colombia: 1	Colombia: 2	Colombia: 11	Colombia: 19	Colombia: 0	Colombia: 0	Colombia: 2
	México: 1	México: 2	México: 25	México: 73	México: 6	México: 4	México: 0
	Venezuela: 1	Venezuela: 2	Venezuela: 24	Venezuela: 0	Venezuela: 69	Venezuela: 2	Venezuela: 2

Fuente: elaboración propia.

Cuadro 5 - Desglose de las características de los artículos científicos procesados usando la metodología de SciELO (Continuación).

Área	Autores	Idiomas	Páginas	Referencias	Notas	Figuras	Tablas
Ciencias Sociales Aplicadas	Brasil: 1	Brasil: 2	Brasil: 30	Brasil: 29	Brasil: 18	Brasil: 0	Brasil: 0
	Chile: 1	Chile: 2	Chile: 33	Chile: 31	Chile: 57	Chile: 0	Chile: 9
	Colombia: 1	Colombia: 2	Colombia: 15	Colombia: 35	Colombia: 0	Colombia: 0	Colombia: 12
	México: 2	México: 4	México: 28	México: 46	México: 0	México: 1	México: 6
	Venezuela: 7	Venezuela: 2	Venezuela: 14	Venezuela: 19	Venezuela: 0	Venezuela: 6	Venezuela: 1
Ingenierías	Brasil: 4	Brasil: 1	Brasil: 16	Brasil: 54	Brasil: 1	Brasil: 5	Brasil: 5
	Chile: 3	Chile: 1	Chile: 8	Chile: 23	Chile: 1	Chile: 0	Chile: 4
	Colombia: 3	Colombia: 3	Colombia: 7	Colombia: 14	Colombia: 1	Colombia: 7	Colombia: 9
	México: 3	México: 1	México: 7	México: 18	México: 1	México: 9	México: 4
	Venezuela: 2	Venezuela: 2	Venezuela: 12	Venezuela: 16	Venezuela: 1	Venezuela: 6	Venezuela: 8
Lingüística, Letras y Artes	Brasil: 1	Brasil: 2	Brasil: 12	Brasil: 25	Brasil: 0	Brasil: 0	Brasil: 3
	Chile: 5	Chile: 2	Chile: 17	Chile: 14	Chile: 6	Chile: 10	Chile: 1
	Colombia: 3	Colombia: 2	Colombia: 33	Colombia: 30	Colombia: 0	Colombia: 9	Colombia: 0
	México: 1	México: 1	México: 16	México: 26	México: 0	México: 3	México: 0
	Venezuela: 1	Venezuela: 2	Venezuela: 30	Venezuela: 33	Venezuela: 1	Venezuela: 1	Venezuela: 3

Fuente: elaboración propia.

En la figura 2 se observa claramente que para Brasil, Chile y México a mayor cantidad de referencias, mayor es el tiempo de procesamiento del artículo científico, mientras que para el caso colombiano y venezolano no se nota esta tendencia, debido a que son pocas las referencias contenidas en los artículos de las diferentes áreas del conocimiento, en comparación con los demás países en estudio.

El comportamiento que se observa en la figura 2 indica que el tiempo total de procesamiento de un artículo científico, usando la metodología de SciELO, se hace mucho mayor por el marcaje detallado de sus referencias y, si se complementa con la información del cuadro 3, se confirma que en promedio el 50.0% de dicho tiempo lo consume la marcación, la cual consiste en el etiquetado de los datos del encabezado (front), el cuerpo (body) y las referencias bibliográficas, agradecimiento, apéndices y el artículo en un idioma distinto del original. De estas tres partes, la que consume la mayor cantidad de tiempo es el back, que contiene la lista de las referencias, debido a que cada una debe ser tratada de forma independiente y detallada.

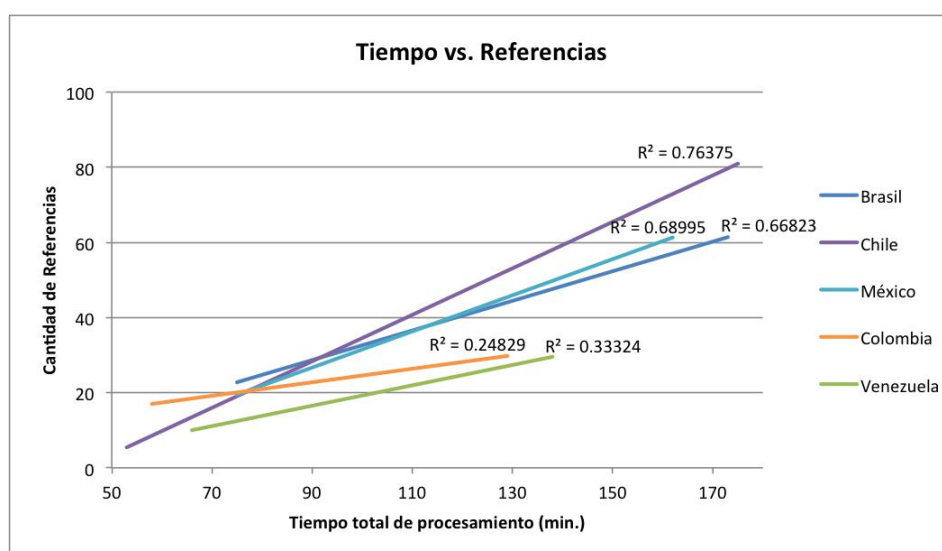


Figura 2 - Comparación de la tendencia del tiempo total de procesamiento con la cantidad de referencias para revistas de Brasil, Chile, México, Colombia y Venezuela

En el cuadro 6 se muestra que las cantidades promedio de referencias por artículo son muy similares para Brasil, Chile y México, siendo aproximadamente el doble que la cantidad de referencias de los artículos de Colombia y Venezuela.

Cuadro 6 - País contra cantidad promedio de referencias de los artículos procesados

País	Cantidad promedio de referencias por artículo	Desviación estándar
Brasil	41	18.7
Chile	36	25.1
Colombia	24	8.3
México	39	18.6
Venezuela	21	11.7

Fuente: elaboración propia.

Respecto a los otros elementos de un artículo científico, como la cantidad de autores, idiomas, figuras, tablas o cuadros, notas o la cantidad de páginas que tenga el artículo, no tienen ninguna influencia significativa en el tiempo total de procesamiento de éste, usando la metodología de SciELO, en ninguno de los países en estudio. En la figura 3 se puede observar la tendencia cuando se contrasta el tiempo total de procesamiento con la cantidad de páginas que tiene el artículo y en la figura 4 se presenta esta misma comparación pero usando la cantidad de tablas y figuras; en este caso, sólo Brasil parece mostrar una relación significativa.

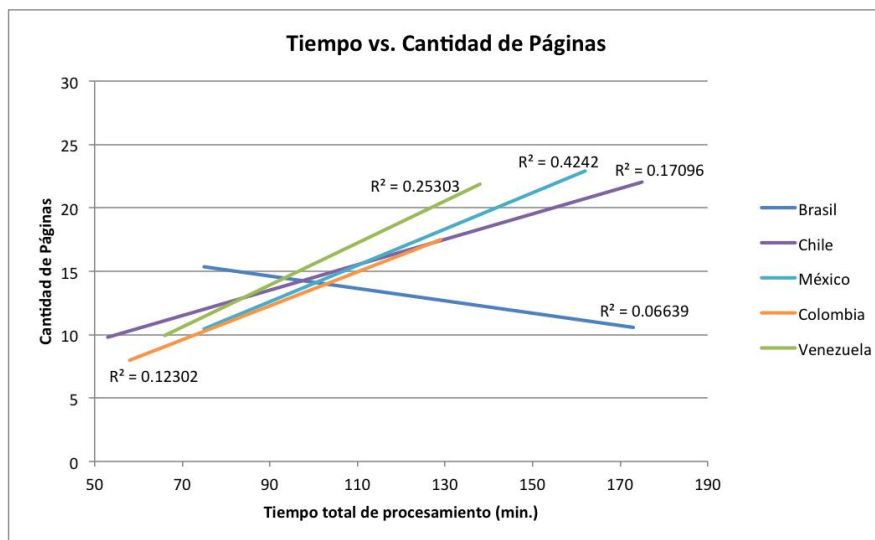


Figura 3 - Comparación de la tendencia del tiempo total de procesamiento con la cantidad de páginas que posee el artículo

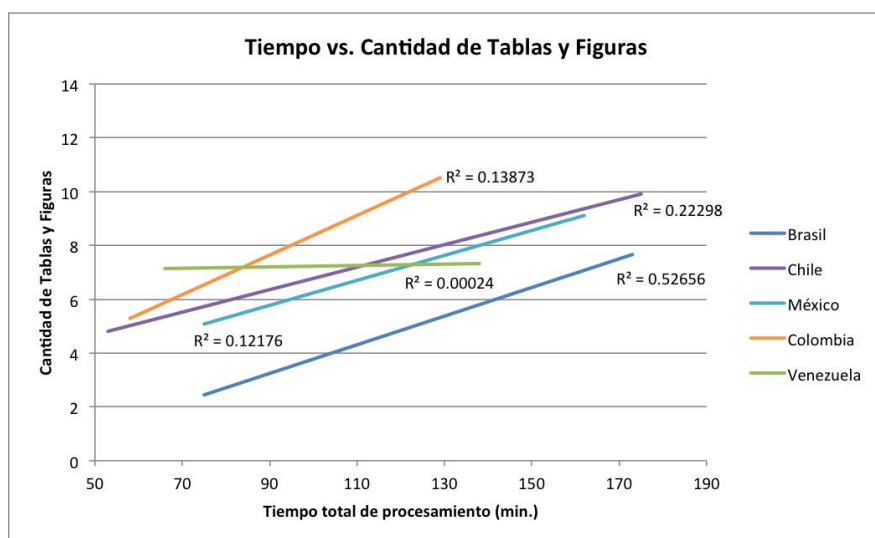


Figura 4 - Comparación de la tendencia del tiempo total de procesamiento con la cantidad de tablas y figuras que posee el artículo

7. Conclusiones

Los sistemas o plataformas de indización latinoamericanas de revistas científicas en acceso abierto han establecido exigencias técnicas a los editores con la finalidad de poder publicar sus contenidos en la web, lo cual ha ocasionado que la mayoría de los editores pierdan el control del proceso editorial y tengan que invertir gran parte de sus presupuestos en generar los contenidos en diferentes formatos, para cumplir las exigencias de los sistemas de indización. Por esta razón, es necesario generar un modelo de comunicación y publicación científica para la región, en el cual el editor se apropie nuevamente del proceso editorial completo, y que por supuesto sea sustentable.

Actualmente, los editores de las revistas científicas tienen el deber de garantizar la preservación digital, la visualización dinámica y la recuperación inteligente del contenido publicado en sus revistas. Esto fácilmente se puede lograr adoptando el uso del estándar JATS, es por esto que no hay discusión de que éste es el estándar por seguir, lo que está en discusión es el cómo se está haciendo. Si se piensa en el modelo latinoamericano de publicación científica en acceso abierto, es importante tomar en cuenta el costo involucrado, debido a que debe ser un modelo sustentable, económica y técnicamente, con la finalidad de que cada país sea capaz de mantenerlo como un proyecto de presupuesto público, de bajo costo para las revistas científicas y sin costo para el usuario final, de tal manera que se evite usar un modelo donde se tercerizan procesos o subcontraten a empresas privadas para generar los archivos XML JATS. Esta forma de trabajo encarece el costo y genera dependencia.

La aplicación de la metodología SciELO a un número completo de una revista científica para generar los archivos XML JATS, según su esquema SPS, es un procedimiento costoso, ya que implica invertir gran cantidad de horas-hombre, debido a que dicha metodología requiere tener experiencia en la marcación, conocimientos técnicos y estar familiarizado con la herramienta (Markup). Por lo anterior, la decisión de usar la metodología y herramientas de publicación de SciELO como modelo de publicación de acceso abierto debe ser pensada y analizada, ya que si no se cuenta con los recursos económicos necesarios se pone en peligro la sustentabilidad del modelo de publicación, así lo demuestran algunos países de la región, por ejemplo Venezuela (scielo.org.ve).

Al emplear el modelo SciELO para la generación de los archivos XML JATS, los tiempos correspondientes al paso 4 (preparación del documento) y al paso 5 (marcación usando el XML SPS) son los que consumen la mayor parte del tiempo de procesamiento de un artículo científico sin importar el área del conocimiento al que pertenezca. El paso 4 tarda en promedio 35.0% ($\sigma = 9.72\%$) del tiempo, mientras que el paso 5 tarda en promedio el 50.0% ($\sigma = 8.50\%$) del tiempo y en conjunto ocupan entre el 70.0% y 90.0% del tiempo total de procesamiento. Analizando esto con mayor detalle, dicho tiempo de procesamiento no se ve influenciado por la cantidad de páginas que tenga el artículo, mientras que la cantidad de referencias bibliográficas sí incrementa este tiempo en gran medida, ya que es la acción que más tiempo toma llevar a cabo por ser un proceso completamente manual y que requiere tener experiencia en la identificación de cada referencia. Respecto a la cantidad de tablas, figuras y ecuaciones que tenga un artículo, se observa que también puede incrementar el tiempo de procesamiento, pero no de forma significativa.

El esfuerzo que lleva a cabo SciELO en la marcación tan detallada de las referencias de un artículo científico tiene mayores frutos si dichos datos se usan en la generación de indicadores bibliométricos basados en la información contenida en las referencias. Sin embargo, no ha sido posible generar indicadores bibliométricos consistentes de citación. SciELO hace uso de la cantidad total de referencias, el año y el título de la revista. Si se quisieran generar únicamente indicadores de citación a nivel de revista, estos dos últimos deberían ser los únicos datos por marcar en una referencia, con la finalidad de reducir los tiempos de procesamiento y en consecuencia los costos de este modelo de publicación.

Por lo expuesto, también es posible que las características técnicas y de diseño de la herramienta de marcación de SciELO estén afectando negativamente los tiempos de procesamiento y en consecuencia su costo. Se debe analizar si una mejora sustancial de dicha herramienta reduce los tiempos de maquilación para obtener el XML JATS, con objeto de garantizar la sustentabilidad del modelo de publicación.

Continuar explorando, basados en evidencias y bajo el principio de replicabilidad, la sustentabilidad de las propuestas latinoamericanas se convierte en factor esencial para decidir los costos, ventajas y límites de la interoperabilidad de los modelos instrumentados en Latinoamérica y poder evaluar la sustentabilidad del modelo de comunicación y publicación científica en acceso abierto de la región.

Referencias

- Alperin, J. P., Fischman, G. E., y Willinsky, J. (2011). Scholarly Communication strategies in Latin America's research-intensive universities. *Revista Educación Superior y Sociedad*, 2(16), 1-9.
- Babini, D. (Agosto de 2012) Biblioteca universitaria: agente de cambio indispensable para comunicaciones académicas y científicas en acceso abierto. En 8va. Jornadas de Bibliotecas y Centros de Documentación de la Universidad de Buenos Aires. Universidad de Buenos Aires, Facultad de Agronomía, Buenos Aires.
- Babini, D., y Machin-Mastromatteo, J. D. (2015). Latin American Science is meant to be open access: Initiatives and current challenges. *Information Development*, 31(5), 477-481. doi: 10.1177/0266666915601420.
- Crow, R., (2013), *Sustainability of Open Access Services. Phase 3: The Collective Provision of Access Resources*, Washington, USA: SPARC.
- Cop, N. (22 de mayo de 2015). ¿Puede el dinero gastado a nivel mundial en suscripciones de revistas ser pasado completamente a un modelo de negocios OA para liberar las revistas? SciELO en Perspectiva. [Mensaje en un blog] Recuperado de: <http://blog.scielo.org/es/2015/05/26/puede-el-dinero-gastado-a-nivel-mundial-en-suscripciones-de-revistas-ser-pasado-completamente-a-un-modelo-de-negocios-oa-para-liberar-las-revistas/>.
- Herrera, L. A., Ferracutti, V. M., Martínez, F. A., y Mitzig, N. L. (2011, octubre 27). El Acceso Abierto para la sustentabilidad de las publicaciones de la Universidad Nacional del Sur. Jornada Virtual Acceso Abierto Argentina 2011. Biblioteca Electrónica de Ciencia y Tecnología - Ministerio de Ciencia, Tecnología e Innovación Productiva / CAICYT-CONICET / OPS-OMS / SPARC. [Archivo de video] Recuperado de <https://www.youtube.com/watch?v=h-lqf9mDqU8>.
- NCBI, National Center for Biotechnology Information, U.S. National Library of Medicine. (2012). Journal Publishing Tag Library NISO JATS Version 1.0. Recuperado de <http://jats.nlm.nih.gov/publishing/tag-library/1.0/index.html>.
- NCBI, National Center for Biotechnology Information, U.S. National Library of Medicine. (2013). JATS: Journal Article Tag Suite. Recuperado de <http://jats.nlm.nih.gov/>.
- Packer, A. L., et al. orgs., (2014), *SciELO – 15 Años de Acceso Abierto: un estudio analítico sobre Acceso Abierto y comunicación científica*. París: UNESCO, 2014, 188 p. ISBN 978-92- 3301-237-0. doi: 10.7476/9789233012370.
- Packer, A. L., Salgado, E., Araujo, J., Aquino, L., Almeida, R., Santos, J., Lucena, S., Soares, C. M. (4 de abril de 2014b). ¿Por qué XML? SciELO en Perspectiva. [Mensaje en un blog] Recuperado de <http://blog.scielo.org/es/2014/04/04/porque-xml/>.
- PKP (Public Knowledge Project). s/f. *Revistas académicas económicamente sustentables y mejor acceso al conocimiento*. Recuperado de https://pkp.sfu.ca/files/sustentabilidad_economica.ppt.
- Rozemblum, C., Bava, L., Unzurrunzaga, C., y Banzato, G. (noviembre de 2014). Costos y beneficios de la inclusión de revistas universitarias en bases de datos de Acceso Abierto. En 12ª Jornada sobre la Biblioteca Digital Universitaria. "La producción científica en la actividad universitaria". Universidad de La Plata, Salta, Argentina. Recuperado de <http://www.unsa.edu.ar/biblio/jbdu/gacetillas/JBDU.pdf>.
- Santos, S. M., y Packer, A. (2014). Producción de la Colección y revistas SciELO. En A. L. Packer, et. al. (Orgs.) *SciELO - 15 Años de Acceso Abierto: un estudio analítico sobre Acceso Abierto y comunicación científica*. (pp. 81-92). París: UNESCO. ISBN 978-92-3301-237-0 Recuperado de <http://dx.doi.org/10.7476/9789233012370>.

SciELO (s/f). Modelo SciELO. Recuperado de

<http://www.scielo.org/php/level.php?lang=es&component=44&item=1>.

SciELO (2013). Grupo de Google SciELO XML - SciELO Publishing Schema. Recuperado de

<https://groups.google.com/forum/#!forum/scielo-xml>.

SciELO. (2014). Indicadores Bibliométricos de la Red SciELO. Recuperado de

<http://www.scielo.org/php/level.php?lang=es&component=44&item=25>

SciELO. (2015a). SciELO Publishing Schema Documentation Versão 1.3. Recuperado de

<https://media.readthedocs.org/pdf/scielo-publishing-schema/latest/scielo-publishing-schema.pdf>

SciELO. (2015b). SciELO PC Programs. Requirements. Recuperado de [http://docs.scielo.org/projects/scielo-pc-](http://docs.scielo.org/projects/scielo-pc-programs/en/latest/howtoinstall.html#requirements)

[programs/en/latest/howtoinstall.html#requirements](http://docs.scielo.org/projects/scielo-pc-programs/en/latest/howtoinstall.html#requirements).

SciELO. (2015c). Preparação de Arquivos para o Programa Markup. Recuperado de

http://docs.scielo.org/projects/scielo-pc-programs/en/latest/pt_how_to_generate_xml-prepara.html#preparacao-de-arquivos-para-o-programa-markup.

SciELO. (2015d). Markup Program. Recuperado de [http://docs.scielo.org/projects/scielo-pc-](http://docs.scielo.org/projects/scielo-pc-programs/en/latest/markup.html)

[programs/en/latest/markup.html](http://docs.scielo.org/projects/scielo-pc-programs/en/latest/markup.html).

SCImago. (2007). SJR - SCImago Journal & Country Rank. Recuperado de <http://www.scimagojr.com>.

Swan, A., (2013), Directrices para políticas de desarrollo y promoción del acceso abierto, UNESCO. ISBN: 978-959-18-0928-5.

Velterop, J. (10 de septiembre de 2015). ¿Qué es lo que está reteniendo la transición al acceso abierto si no cuesta más? SciELO en Perspectiva. [Mensaje en un blog] Recuperado de

<http://blog.scielo.org/es/2015/09/10/que-es-lo-que-esta-reteniendo-la-transicion-al-acceso-abierto-si-no-cuesta-mas/>.

Datos de los autores

Eliana Guzmán-Useche

Docente de la Facultad de Ingeniería de la Universidad de Los Andes, Mérida, Venezuela.

eliguz28@gmail.com

Fernando Rodríguez-Contreras

Coordinador del Índice de Revistas Venezolanas de Ciencia y Tecnología en los Servicios Bibliotecarios de la Universidad de Los Andes, Mérida, Venezuela.

frodrig@ula.ve

Recibido - Received: 2016-03-01

Aceptado - Accepted: 2016-05-15



This work is licensed under a Creative Commons Attribution 4.0 United States License.



This journal is published by the [University Library System](#) of the [University of Pittsburgh](#) as part of its [D-Scribe Digital Publishing Program](#) and is cosponsored by the [University of Pittsburgh Press](#).