

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

Eliana Guzmán-Useche

Fernando Rodríguez-Contreras

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

ORIGINAL

Resumen

Objetivo: realizar el análisis de sustentabilidad de la iniciativa latinoamericana Redalyc.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 Marcalyc. Se procesan 40 artículos publicados en revistas científicas de diferentes áreas del saber, pertenecientes a Brasil, Chile, Colombia, México y Venezuela, usando dicha herramienta.

Resultados: se obtiene que la marcación tarda en promedio el 86,3% (SD=4,2%) del tiempo total de procesamiento, repartido en partes casi iguales entre el contenido del artículo (*body*) (media=48,3%, SD=9,5%) y las referencias, notas, agradecimiento, apéndice y glosario (*back*) (media=43,2%, SD=9,3%).

Conclusiones: La sustentabilidad del modelo de publicación de Redalyc, está más garantizada en el aspecto económico por dos razones, los tiempo de procesamiento para generar el XML-JATS son mucho menores que los de SciELO y además hace uso de una herramienta de marcación (Marcalyc), que puede ser utilizada directamente por el equipo editorial de una revista, sin necesidad de conocimientos técnicos especializados. Esto favorece la sustentabilidad del modelo de publicación científica en acceso abierto latinoamericano y devuelve el control del proceso de publicación a los editores.

Palabras clave

Acceso abierto. JATS. Latinoamérica. Marcalyc. Publicación científica. Redalyc. Revista científica. SciELO. Sustentabilidad. XML.

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

Abstract

Objective. To carry out the sustainability analysis of the Latin American initiative Redalyc.org, based on two fundamental aspects, the economic and the need for its existence.

Method. the measurement of the man-hours required to publish a scientific article is carried out, using Marcalyc. 40 articles published in scientific journals from different areas of knowledge, belonging to Brazil, Chile, Colombia, Mexico and Venezuela, are processed using this tool.

Results. it is obtained that the markup takes on average 86.3% (SD = 4.2%) of the total processing time, distributed in almost equal parts between the content of the article (*body*) (average = 48.3%, SD = 9.5%) and references, notes, acknowledgement, appendix and glossary (*back*) (mean = 43.2%, SD = 9.3%).

Conclusions. The sustainability of Redalyc's publication model is more economically guaranteed for two reasons; 1) the processing time to generate XML-JATS is much shorter than SciELO, and 2) it uses a markup tool (Marcalyc), which can be used directly by the editorial team of a journal, without the need for specialized technical knowledge. This favors the sustainability of the scientific publication model in Latin American open access and returns the control of the publication process to the editors.

Keyword

JATS. Latin America. Marcalyc. Open access. Redalyc. SciELO. Scientific journal. Scientific publication. Sustainability. XML.

1 Introducción

El presente trabajo es la continuación de la investigación que se lleva a cabo entre la Universidad de Los Andes (Venezuela) y la Universidad Autónoma del Estado de México (México), sobre “Acceso Abierto y Sustentabilidad en América Latina”, cuyo objetivo es evaluar el aspecto técnico de los modelos de publicación y comunicación de la ciencia, implementados por los dos principales actores de la región: scielo.org (CNPq-Brasil) y redalyc.org (UAEM-México).

El acceso abierto es una filosofía que busca reforzar la difusión de los resultados de la actividad científica. En Latinoamérica ha sido de gran ayuda para la comunidad científica y académica, ya que en conjunto con las tecnologías de la información y comunicación (TIC), han permitido realizar dicha difusión de una forma mucho más eficiente y al alcance de estudiantes, profesores, investigadores y todo aquel que esté interesado en consultar contenido científico y riguroso de forma gratuita. De allí que sea importante diseñar un modelo latinoamericano de publicación de contenido científico en Acceso Abierto (AA), el cual debe ser sustentable para todos los países de la región y a la vez hacer uso de los avances tecnológicos disponibles.

En un primer trabajo (Guzmán y Rodríguez, 2016), se realizó la medición de las horas-hombre requeridas para obtener el formato XML-JATS del contenido de una revista científica usando la metodología de SciELO, con la finalidad de evaluar la sustentabilidad de dicho modelo, desde una perspectiva técnica. Dicho trabajo permitió concluir que la aplicación de la metodología de SciELO es costosa, porque consume muchas horas-hombre, el proceso de aprendizaje es largo y requiere poseer bastantes conocimientos técnicos sobre *eXtensible Markup Language* (XML) y *HyperText Markup Language* (HTML). Por lo tanto, dicha metodología no es usada directamente por los editores de las revistas, quienes se ven obligados a contratar empresas para cumplir con los requerimientos de la generación del contenido para los sitios SciELO.

Lo antes expuesto implica que la sustentabilidad del modelo de Acceso Abierto (AA) en América Latina no estaría garantizada, si la decisión es seguir el modelo de generación y publicación de contenidos de SciELO, ya que dicha sustentabilidad se garantiza, según Crow (2013), por dos aspectos claves: la necesidad de la existencia de la iniciativa y el aspecto económico. El primero de estos aspectos no estaría en duda, pero en cuanto al económico ha quedado demostrado que la propuesta de SciELO es costosa (Guzmán y Rodríguez, 2016).

El objetivo de este trabajo es evaluar la propuesta metodológica para la generación del contenido de una revista científica de AA en formato XML-JATS, realizada por Redalyc, usando la misma muestra de artículos y revistas que se procesó en la evaluación de la metodología de SciELO, con la finalidad de que se puedan comparar estos dos modelos y aportar los datos necesarios para iniciar o continuar con la discusión sobre el modelo de publicación en AA de la región.

Los aspectos a considerar en la evaluación antes mencionada contemplan las características de la herramienta de marcación (Marcalyc), las recomendaciones hechas por Redalyc en la forma cómo se deben marcar los contenidos y en la interoperabilidad del XML-JATS generado.

2 Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc)

La Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc) fue pensada y diseñada por un grupo de investigadores de la Universidad Autónoma del Estado de México (UAEM) en el año 2002, con el objetivo de sortear los obstáculos para el acceso a la información científica producida en el mundo, además de mejorar la escasa visibilidad de la información científica producida en la región.

En el año 2003 inicia como una hemeroteca digital que indiza a un conjunto de revistas iberoamericanas en acceso abierto, a través del portal en línea de Redalyc, <http://www.redalyc.org>. Comenzó publicando el contenido de 28 revistas de ciencias sociales y humanidades, pero a partir del año 2006 comenzaron a ingresar revistas de otras áreas del conocimiento (Revista Digital Universitaria, 2014). Esta cantidad inicial de revistas se ha ido incrementando tal y como se muestra en la Figura 1, donde se aprecia que la actividad científica editorial de la región presentaba grandes necesidades que fueron cubiertas por los objetivos iniciales de Redalyc, de ahí su apoyo y el querer formar parte de él.

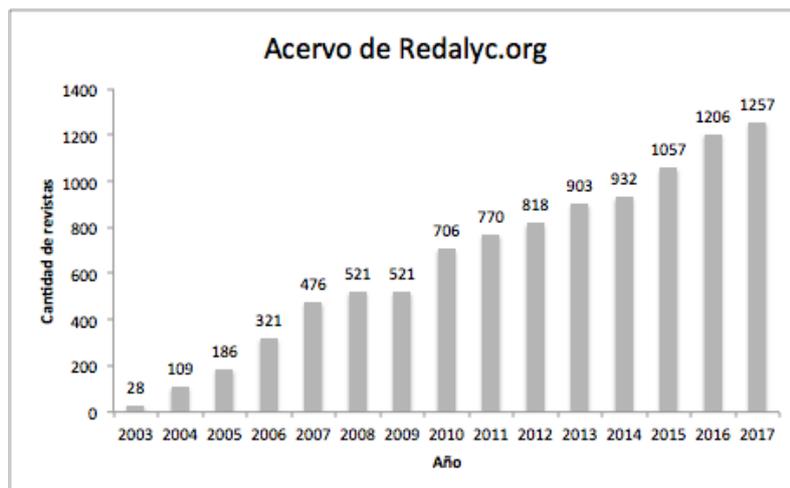


Figura 1 - Evolución de la cantidad de revistas indizadas en Redalyc

Para finales de 2017 Redalyc indiza 1257 revistas, aunque tiene aceptadas 1350 revistas que cumplen con los criterios editoriales solicitados, pero según los nuevos acuerdos de indización se publicarán en Redalyc.org una vez que comiencen a enviar su contenido en formato XML bajo el estándar JATS (Aguado y Becerril, 2017).

En el año 2008 Redalyc se constituye como un sistema de información científica al diseñar indicadores que permiten analizar la producción científica, su difusión y uso tomando como muestra los datos de las revistas indizadas, además de generar informes que fueron producto de dichos análisis. Actualmente el trabajo de Redalyc se enfoca en cuatro aspectos, continuar generando el contenido de la hemeroteca científica, diseñando indicadores cuantitativos, produciendo informes de producción científica, ya sea por áreas del conocimiento o por países y realizando los aportes necesarios para construir un nuevo modelo de publicación científica en acceso abierto, no comercial y sustentable (Olvera-García et. al., 2015).

Desde su inicio hasta la presente fecha, Redalyc publica el contenido de los artículos científicos, usando el formato *Portable Document Format* (PDF), una vez que se han registrado los metadatos de cada artículo en una base de datos. En el año 2015, Redalyc anuncia que hará uso del XML bajo el estándar *Journal Article Tag Suite* (JATS) como formato primario para la publicación y preservación de los contenidos de los artículos. En agosto del año 2016, libera la herramienta de marcación, Marcalyc, dirigida a los editores de las revistas, con la finalidad de que puedan marcar cada artículo y generar el formato XML-JATS, del cual se obtienen otros formatos como PDF, ePub, HTML enriquecido, XML según el esquema de publicación de SciELO v. 1.6 y dos visores, uno para los lectores que usen una computadora y otro para quienes usen un teléfono inteligente o tableta.

3 Materiales y método

3.1. Selección de la muestra de revistas

La muestra de revistas científicas empleada en este artículo es la misma utilizada por Guzmán y Rodríguez (2016), con la finalidad de medir y posteriormente comparar los tiempos de marcación al usar dos herramientas para la generación de los archivos XML bajo el estándar JATS, con las que cuentan los editores de las revistas científicas de América Latina y el Caribe. Por tal motivo, se mantiene la misma clasificación de los artículos científicos según su tipo (editorial, artículo original, reseña, ensayo, avance de investigación, entre otros) y área de conocimiento, tomando en cuenta que el contenido y extensión puede ser muy diferente.

La muestra está formada por 40 artículos publicados en revistas científicas de Brasil, Chile, Colombia, México y Venezuela como países que representan diferentes circunstancias en cuanto a la producción científica en Latinoamérica. La selección se hizo tomando en cuenta el ranking de revistas del Scimago Journal & Country Rank de 2014 (SCImago, 2007), escogiendo la revista, por área de conocimiento, que tenga el mayor SJR de su

país de edición. Los detalles de cada uno de los artículos que forman parte de la muestra se encuentran en el cuadro 1 de Guzmán y Rodríguez (2016).

Se mantiene la misma clasificación de las áreas de conocimiento de las revistas, la cual contempla cuatro áreas que pertenecen a las ciencias sociales y humanas, mientras que las otras cuatro son de las ciencias naturales y exactas:

- 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.

3.2. Materiales

En el procesamiento de los artículos de la muestra se usa Marcalyc v. 1.5, y se sigue el procedimiento indicado en el manual de usuario (Aguado-López, et-al, 2016) y los otros recursos de ayuda (videos y preguntas frecuentes) disponibles en <http://marcalyc.redalyc.org/ayuda/>.

Para generar los archivos HTML se utiliza Microsoft Word® 2010 y para los archivos ZIP se utiliza WinRAR 5.10 (32 bit).

Con respecto al recurso humano que lleva a cabo el experimento cuenta con más de 10 años de experiencia en la marcación de artículos usando la herramienta de SciELO (Markup) y durante un mes se capacitaron en el uso de Marcalyc para familiarizarse con la herramienta y luego si procesar la muestra de los 40 artículos seleccionados.

4 Resultados

Los principales resultados de este experimento son los tiempos de procesamiento de un artículo científico para obtener su versión final en formato XML bajo el estándar JATS, usando Marcalyc. En dicho procesamiento se pueden distinguir los siguientes pasos:

- **Paso 1.** A partir de la versión final del artículo se genera el archivo index.html usando la opción de Microsoft Word® “Guardar como” y verificando la codificación UTF-8 en las “Opciones web...”. Posteriormente, se comprime este archivo junto con la carpeta de archivos adicionales (index_archivos), para generar el archivo ZIP que requiere Marcalyc para comenzar la marcación.
- **Paso 2.** Crear, modificar o verificar los archivos gráficos que pueda contener el artículo. Por razones de visualización en los productos de la marcación que genera Marcalyc a partir del XML, los archivos gráficos deben cumplir con ciertos requisitos de dimensiones de las imágenes en pixeles, además el tamaño de cada archivo no debe ser mayor a 1 MB.
- **Paso 3.** Se debe crear el número de la revista con el que se va a trabajar. Este paso se realiza una sola vez para todos los artículos.
- **Paso 4.** Agregar el artículo que se va a marcar a la tabla de contenido del número.
- **Paso 5.** Marcación usando el Marcalyc.
- **Paso 6.** Depuración de errores e incongruencias de la marcación: cuando se finaliza la marcación y se procede a generar el archivo XML es posible que existan errores, lo que requiere revisar los mensajes de error que se generan para proceder a corregirlos, en el caso de las incongruencias se deben revisar y decidir si es posible corregirlas o no.

- **Paso 7.** Generación del XML y sus productos: una vez que se han revisado las incongruencias y corregidos los errores, se procede a generar el archivo XML y sus productos: PDF, ePub, visor para computadoras, visor para móviles y tabletas, HTML enriquecido y XML según el SPS de SciELO.

4.1. Medición de tiempos

En la Figura 2 se presentan el tiempo total de procesamiento utilizando Marcalyc para generar los archivos XML JATS. Si se discrimina dicho tiempo en cada uno de los pasos antes mencionados, se observa que:

- Generar el archivo index.html y comprimirlo, toma entre 30 segundos y 1 minuto.
- Crear, modificar o verificar los archivos gráficos toma entre 1 y 9 minutos.
- Crear el número a ser marcado, en este experimento se hizo una sola vez, tarda 1,75 minutos.
- Agregar el artículo a la tabla de contenido toma entre 1 y 2 minutos.
- La marcación es el proceso que más tiempo ocupa, entre 25 y 111 minutos.
- Depurar de errores e incongruencias toma entre 0 (si no hay ningún error) y 6 minutos.
- Generar el archivo XML al finalizar el marcado toma entre 6 segundos y 5 minutos.

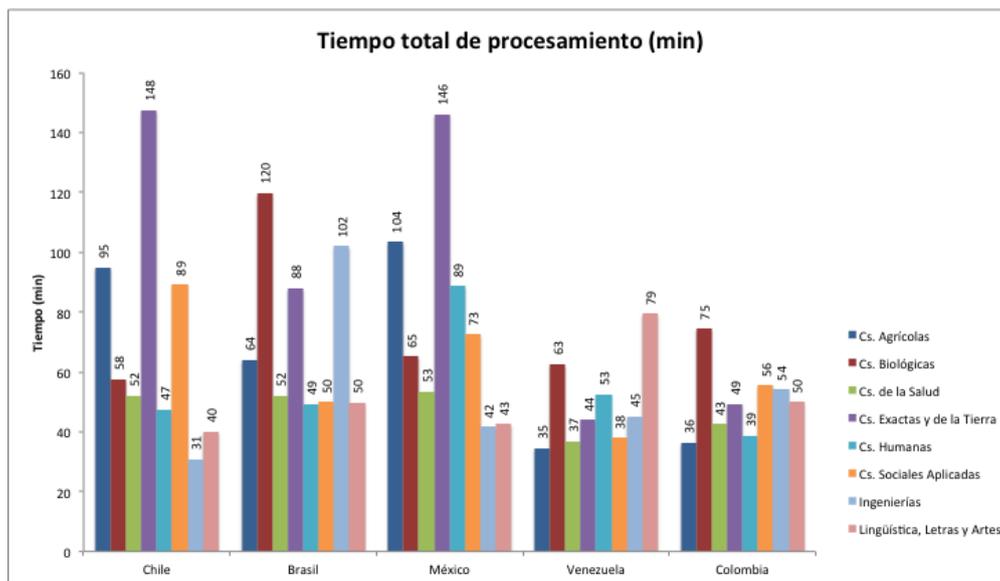


Figura 2 - Tiempo total de procesamiento usando Marcalyc

5 Análisis

5.1. Herramienta de Marcación de Redalyc: Marcalyc

Marcalyc es una herramienta de marcación de artículos científicos que permite obtener el formato XML-JATS y a partir de éste se generan diferentes productos de la marcación como visores (PC y móvil), PDF, HTML enriquecido y ePub. Además, ofrece la ventaja de poder descargar el archivo XML (denominado XML-JAT4R) y el XML-SciELO que sigue el esquema de publicación de SciELO (SPS v. 1.6), lo cual garantiza que la revista pueda utilizar dichos archivos como mecanismo más confiable de preservación de sus contenidos.

Al evaluar a Marcalyc desde el punto de vista técnico como usuario, se observan las siguientes características:

1. Funciona en ambiente web, lo que implica que se puede usar en múltiples plataformas como Linux, MAC OS o Windows y hace uso de las nuevas tecnologías disponibles. Esta forma de trabajo puede ser una desventaja para países con problemas de acceso a Internet o de velocidad de conexión.
2. Todos los elementos de Marcalyc (formularios, botones y listas de valores) se encuentran en español y no hace uso de tecnicismos, por lo tanto, el tiempo de capacitación en el uso de esta herramienta es corto.
3. El usuario de Marcalyc no requiere tener conocimientos sobre las etiquetas y atributos del XML o del estándar JAST, debido a que las advertencias y errores en la marcación se corrigen usando exclusivamente la herramienta. No requiere de ningún tipo de editor adicional, ya que está garantizado que el XML obtenido pasa los validadores de Pub Med Central (PMC XML Validator: <https://www.ncbi.nlm.nih.gov/pmc/tools/xmlchecker>) y de JATS for Reuse (JATS4R: <http://jats4r.org/validator>).
4. Marcalyc realiza un reconocimiento automático de ciertos datos del artículo, en el Front infiere los resúmenes y palabras claves, mientras que en el Back infiere las notas y en las referencias los autores y el año. Lo cual implica que se reducen los tiempos de marcación.
5. No es necesario transcribir, ya que la marcación usando Marcalyc se basa en la selección y pegado automático del texto, reduciendo los errores que se pudieran generar.
6. Requiere tener una cuenta de usuario asignada por Redalyc, no es una herramienta disponible para cualquier revista que no quiera utilizar, ya que sólo puede ser usada por revistas que están indizadas en Redalyc.

En general, Marcalyc es una herramienta bastante intuitiva para un usuario que esté familiarizado con la estructura de un artículo científico, y algo importante de resaltar es que el usuario no necesita poseer conocimientos técnicos sobre XML, sus etiquetas y atributos, ya que es muy sencillo resolver los errores que se puedan cometer durante la marcación.

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

Cuando se comparan los tiempos medidos en cada uno de los siete pasos antes explicados y que son necesarios para obtener el XML de un artículo usando Marcalyc, se observa que el paso 5, que corresponde a la marcación, es el que consume la mayor parte del tiempo total de procesamiento del artículo sin importar su área de conocimiento ni país al que pertenezca la revista, con un promedio de 86,2% ($\sigma=4,2\%$) del tiempo. En el Cuadro 1 se presentan los porcentajes del tiempo de marcación usando Marcalyc, para los diferentes países y áreas de conocimiento, donde se observa que son valores muy similares ya que están poco dispersos.

Cuadro 1. Porcentajes de tiempo de marcación (paso 7) usando Marcalyc.

Área	Brasil	Chile	Colombia	México	Venezuela
Ciencias Agrícolas	88,2	84,2	83,5	88,9	82,5
Ciencias Biológicas	92,7	86,7	93,1	90,4	81,5
Ciencias de la Salud	78,8	80,7	90,2	88,6	84,4
Ciencias Exactas y de la Tierra	89,3	92,8	77,7	86,1	83,5
Ciencias Humanas	86,7	84,8	81,8	92,7	87,5
Ciencias Sociales Aplicadas	92,0	81,8	86,5	84,6	82,7
Ingenierías	85,8	83,2	87,3	82,2	84,8
Lingüística, Letras y Artes	89,7	86,5	89,8	81,2	92,8
Promedio	87,9	85,1	86,2	86,8	85,0
Desviación estándar	4,4	3,8	5,0	4,0	3,7

Fuente: elaboración propia.

El porcentaje faltante de los tiempos de procesamiento se reparten casi de forma igualitaria en los otros seis pasos antes mencionados. Una de las ventajas de usar Marcalyc es que no requiere una preparación previa del documento con un formato en particular, sólo se debe guardar la versión final del artículo como HTML y comprimir en un archivo ZIP.

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

En el Cuadro 2 se presentan los tiempos promedio de procesamiento de un artículo científico usando Marcalyc, para generar su correspondiente archivo XML JATS, según su área del conocimiento.

Cuadro 2. Tiempo promedio de procesamiento de un artículo científico.

Área	Mediana (min)	Promedio (min)	Desviación estándar (min)
Ciencias Agrícolas	64,1	66,7	32,1
Ciencias Biológicas	65,3	75,9	25,2
Ciencias de la Salud	52,1	47,4	7,4
Ciencias Exactas y de la Tierra	88,0	95,0	50,2
Ciencias Humanas	49,3	55,3	19,3
Ciencias Sociales Aplicadas	55,5	61,1	20,1
Ingenierías	45,2	54,8	27,6
Lingüística, Letras y Artes	49,7	52,4	15,8

Fuente: elaboración propia.

A partir de estos tiempos medidos se calcula el tiempo estimado para procesar un número completo de una revista, el cual es proporcional a la cantidad de artículos publicados en dicho número. En el Cuadro 3 se muestran estos tiempos, según el área de conocimiento de la revista, considerando el ejemplo de un número que publica doce artículos. Se puede observar que estos tiempos promedio varían entre 9,5 horas ($\sigma=1,5$ h) para Ciencias de la Salud y 19 horas ($\sigma=10,0$ h) para Ciencias Exactas y de la Tierra.

Cuadro 3. Tiempo promedio de procesamiento de un número publicado por una revista, según su área de conocimiento.

Área	Tiempo promedio (h)	Desviación estándar (h)
Ciencias Agrícolas	13,3	6,4
Ciencias Biológicas	15,2	5,0
Ciencias de la Salud	9,5	1,5
Ciencias Exactas y de la Tierra	19,0	10,0
Ciencias Humanas	11,1	3,9
Ciencias Sociales Aplicadas	12,2	4,0
Ingenierías	11,0	5,5
Lingüística, Letras y Artes	10,5	3,2

Fuente: elaboración propia.

5.4. Tiempo total de procesamiento y su relación con el contenido del artículo

Al analizar la relación existente entre el tiempo total de procesamiento (TTP) de un artículo usando Marcalyc y su contenido, se puede apreciar que no existe una relación significativa con respecto a la cantidad de caracteres, figuras, tablas, ecuaciones, páginas y notas al pie de página que formen parte del artículo.

Con respecto a la relación existente entre el TTP y la cantidad de referencias, se mantiene el mismo comportamiento observado en los trabajos de Herrera et. al (2011) y Guzmán y Rodríguez (2016). Como se

observa en la Figura 3, para Brasil, Chile y México a mayor cantidad de referencias mayor es el TTP, mientras que para Colombia y Venezuela no existe tal comportamiento.

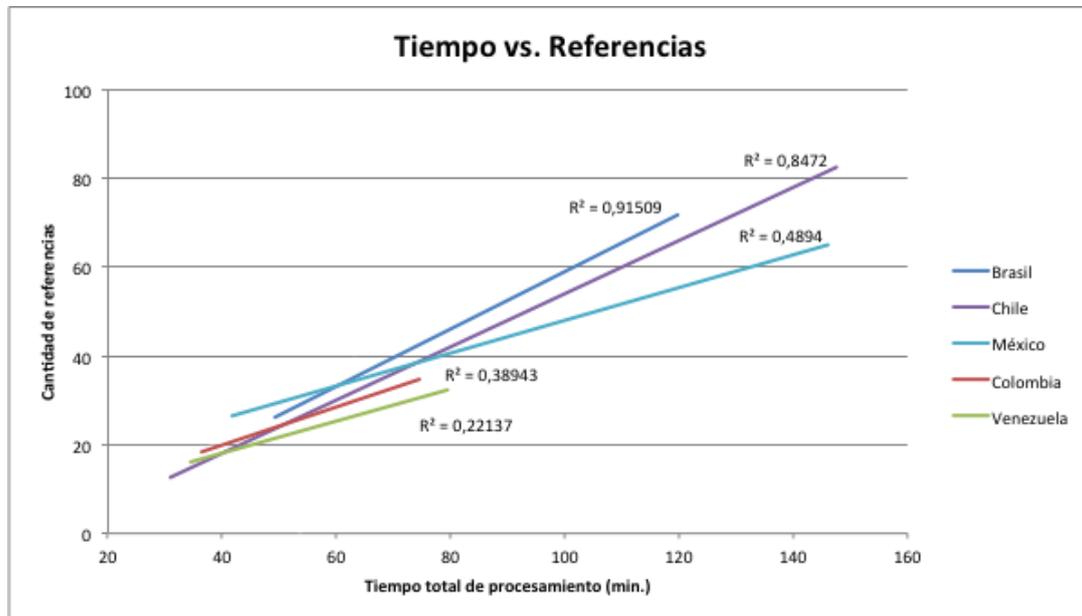


Figura 3 - Comparación de la tendencia del tiempo total de procesamiento con respecto a la cantidad de referencias para revistas de Brasil, Chile, México, Colombia y Venezuela

5.5. Comparación de los tiempos de procesamiento usando la herramienta de marcación de SciELO (Markup v. 4.0.090-trial) y usando la herramienta de Redalyc (Marcalyc v. 1.5)

En el Cuadro 4 se presentan los tiempos totales de procesamiento de la muestra de artículos, usando las metodologías de SciELO y Redalyc para generar el formato XML-JATS. Para el paso específico de marcación, en el caso de SciELO se utiliza la herramienta Markup, y para Redalyc se utiliza Marcalyc. Como se puede observar en todos los artículos el tiempo total de procesamiento utilizando Marcalyc es menor que el tiempo usando Markup, para todas las áreas del conocimiento y países.

Cuadro 4. Tiempos de procesamiento por artículo en minutos usando el Markup de SciELO y Marcalyc de Redalyc.

		Brasil	Chile	Colombia	México	Venezuela
Ciencias Agrícolas	SciELO	148	97	80	131	126
	Redalyc	64	95	36	104	35
Ciencias Biológicas	SciELO	168	89	129	106	136
	Redalyc	120	58	75	65	63
Ciencias de la Salud	SciELO	105	107	117	89	93
	Redalyc	52	52	43	53	37
Ciencias Exactas y de la Tierra	SciELO	124	175	85	162	114
	Redalyc	88	148	49	146	44
Ciencias Humanas	SciELO	79	112	58	139	96
	Redalyc	49	47	39	89	53
Ciencias Sociales Aplicadas	SciELO	75	110	99	145	66
	Redalyc	50	89	56	73	38
Ingenierías	SciELO	173	53	106	84	94
	Redalyc	102	31	54	42	45
Lingüística, Letras y Artes	SciELO	150	76	112	75	138
	Redalyc	50	40	50	43	79

Fuente: elaboración propia.

Al analizar de forma más detallada los tiempos promedio de procesamiento se puede observar que la mayor diferencia se observa en las áreas de ciencias agrícolas, ciencias de la salud, ciencias humanas, ingenierías y lingüística, letras y artes (ver Figura 4), donde las diferencias se aproximan al 50% del tiempo, lo que implica que el tiempo de procesamiento usando Marcalyc se pueden reducir a la mitad, con respecto al tiempo empleado usando la metodología de SciELO (ver Figura 5).

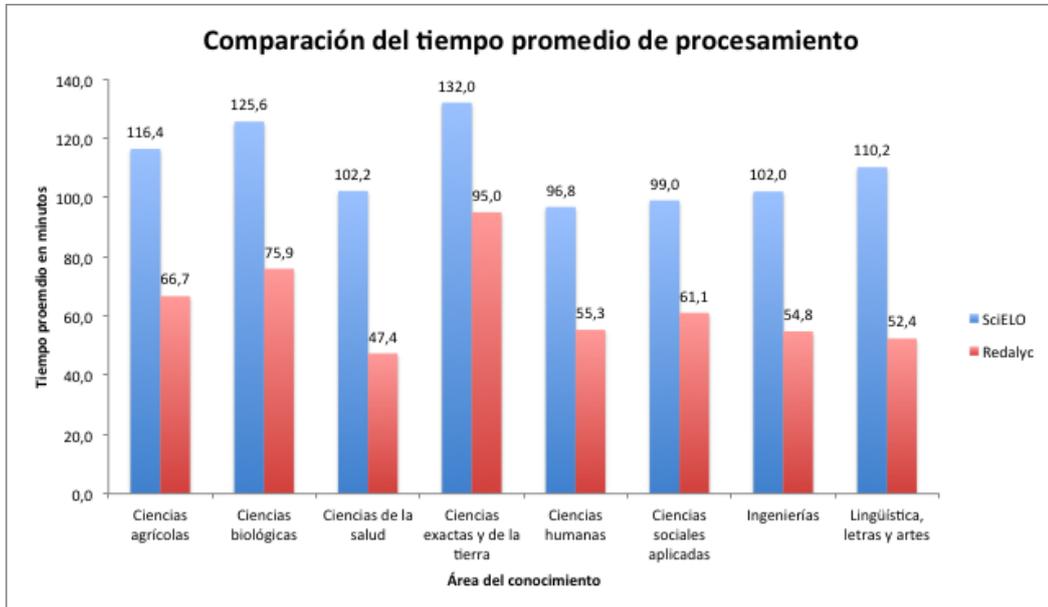


Figura 4 - Comparación del tiempo promedio de procesamiento entre SciELO y Redalyc, según el área de conocimiento

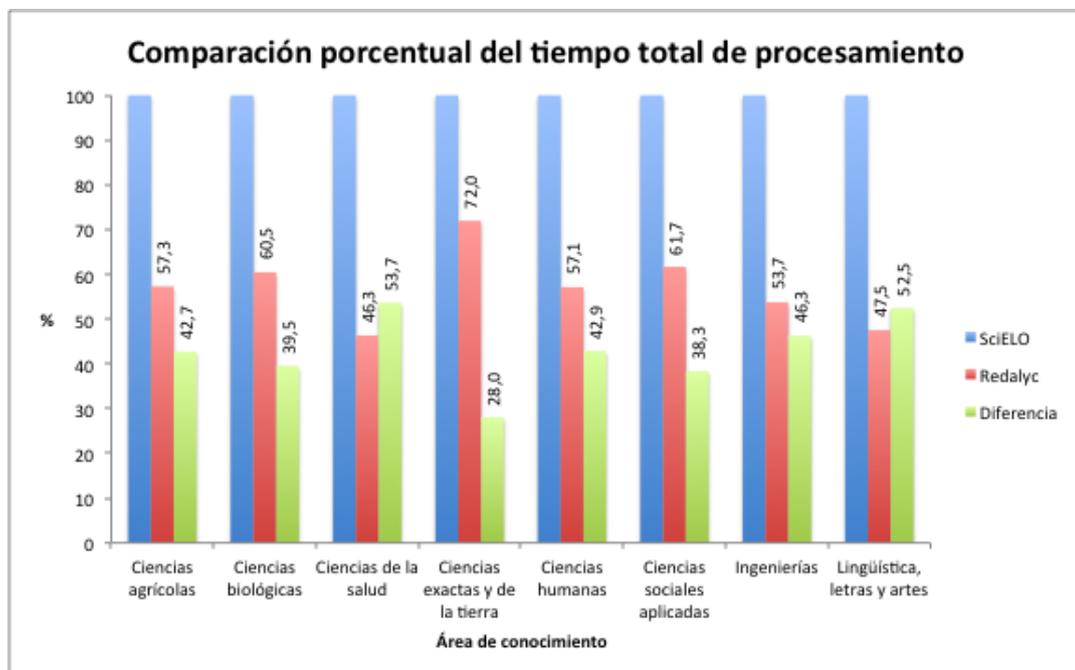


Figura 5 - Diferencia porcentual del tiempo promedio de procesamiento entre SciELO y Redalyc, según el área de conocimiento

Además, si se analizan los tiempos promedio de procesamiento según el país de adscripción de la revista, se observa que de nuevo estos tiempos son menores al usar Marcalyc (ver Figura 6). La diferencia porcentual está entre el 32% al 55%, como se puede ver en la Figura 7.

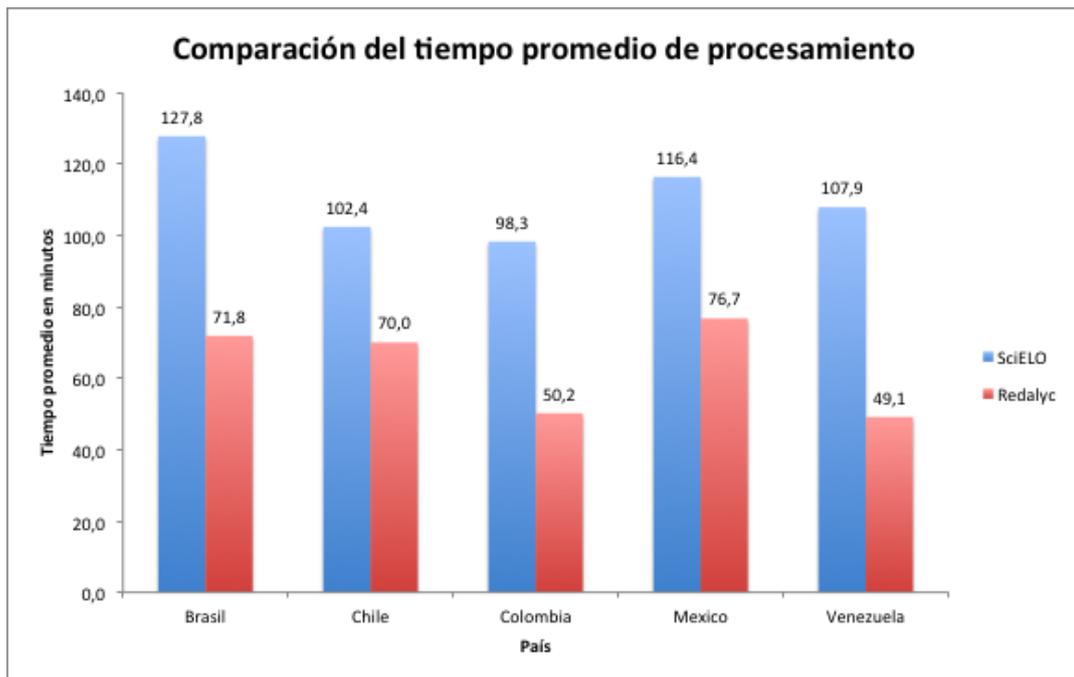


Figura 6 - Comparación del tiempo promedio de procesamiento entre SciELO y Redalyc, por país

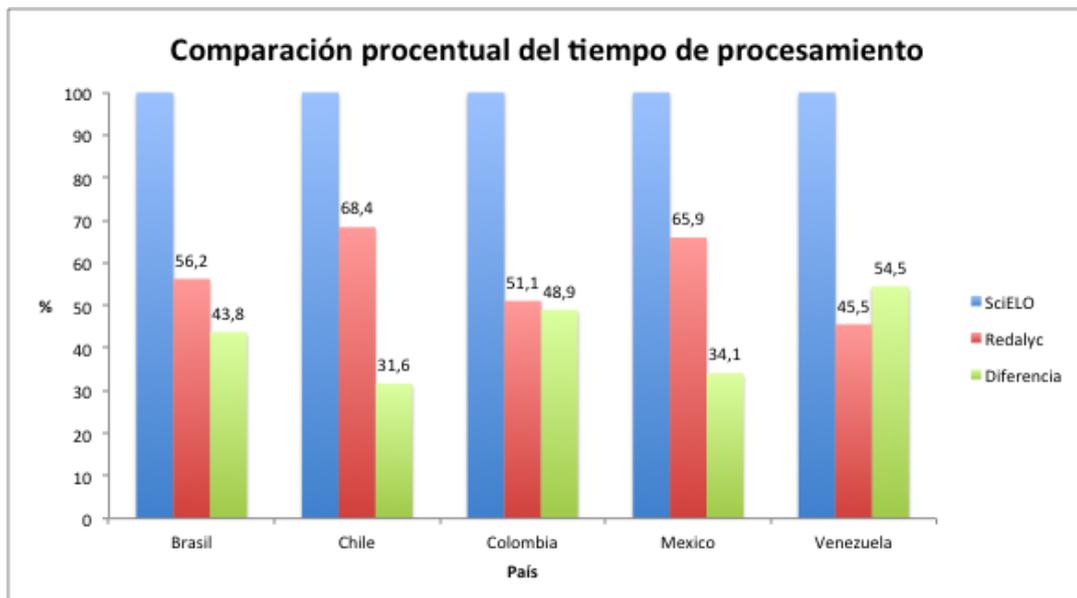


Figura 7 - Diferencia porcentual del tiempo promedio de procesamiento entre SciELO y Redalyc, por país

5 Conclusiones

El modelo de comunicación y publicación científica en acceso abierto de Latinoamérica debe tomar en cuenta las condiciones particulares de las revistas científicas editadas en la región, con el fin de establecer los mecanismos apropiados que garanticen su sustentabilidad, desde el punto de vista económico y técnico. Dicho modelo ha comenzado a tomar forma en las propuestas planteadas por los dos principales metaeditores de la región, Redalyc y SciELO, al establecer criterios de calidad editorial para las revistas, dar visibilidad a los contenidos, diseñar índices bibliométricos, capacitar a los editores en sus funciones, garantizar la preservación de los contenidos, diseñar herramientas que soporten la actividad editorial, entre otras.

Para llevar a cabo gran parte de estas tareas, tanto SciELO como Redalyc, han adoptado el XML bajo el estándar JATS como el formato primario del cual se generan otros formatos de publicación de un artículo científico, por lo tanto, los editores de las revistas deben ser capaces de producir la versión final de sus contenidos en XML-JATS. Existen diferentes herramientas que pueden usarse para generar un archivo XML-JATS, siendo algunas muy costosas y con grandes exigencias de conocimientos técnicos que no necesariamente posee el editor de una revista. Por esta razón, Redalyc y SciELO han puesto a disposición de los editores herramientas para realizar la marcación de un artículo científico y así obtener el archivo XML bajo el estándar JATS.

Al analizar Marcalyc, que es la herramienta diseñada por Redalyc para que el editor genere el formato XML-JATS del contenido de su revista, se puede destacar que su uso no requiere poseer conocimientos técnicos sobre formatos de publicación electrónica, lo que la hace más amigable e intuitiva para los editores de las revistas que no están familiarizados con dichos conocimientos. Este es un aspecto muy importante, ya que empodera al editor y lo hace apropiarse de todos los procesos editoriales de la revista, al no verse obligado a contratar personal especializado para realizar esta labor que ahora le solicitan, tanto Redalyc como SciELO.

Con respecto a los tiempos de procesamiento de un artículo usando Marcalyc, se observa que son significativamente menores que los tiempos de procesamiento usando Markup (herramienta de SciELO), ya sea que se analicen por área del conocimiento o por país de edición de la revista. Al igual que para SciELO, al usar Marcalyc el paso que corresponde a la marcación es el que más tiempo consume, en promedio 86,2% ($\sigma=4\%$). Pero cuando se analiza el tiempo total de procesamiento con Marcalyc, se observa que es significativamente menor que al usar el Markup, básicamente por dos razones, no es necesario darle formato al contenido del artículo para iniciar la marcación y la corrección de errores es mucho más rápida y sencilla. Esta diferencia en el tiempo total de procesamiento es en promedio 43% ($\sigma=8\%$) menor al usar Marcalyc.

Cuando se analiza el contenido del artículo para determinar su influencia en el tiempo total de procesamiento, se observa que a mayor cantidad de referencias mayor es dicho tiempo. Es necesario resaltar que se marcaron todos los datos de cada referencia, tal y como lo establece la metodología de marcación de SciELO, ya que, por su parte, Redalyc, sólo exige que se marquen los autores, año y fuente de cada referencia. Al realizar la marcación de esta manera es lógico que el tiempo sea aún mucho menor. Sería necesario establecer, la utilidad de realizar la marcación de todos los datos de una referencia, ya que aumentan el tiempo total de procesamiento haciéndolo más costoso. Con respecto a otros contenidos del artículo como tablas, ecuaciones, cantidad de páginas, etc., no tienen una influencia significativa en el tiempo total de procesamiento, cuando se utiliza Marcalyc.

A partir de lo antes expuesto, se confirma que las características técnicas de la herramienta de SciELO (Markup), incrementan en gran medida los tiempos totales de procesamiento de un artículo científico, lo cual hace que no se pueda garantizar la sustentabilidad económica de un modelo de comunicación y publicación científica en acceso abierto de la región que haga uso de esta herramienta.

Otro aspecto importante es la usabilidad de estas dos herramientas (Markup de SciELO y Marcalyc de Redalyc), se debe señalar que ambas son gratuitas, aunque Marcalyc sólo está disponible para editores de revistas que estén indizadas en Redalyc, mientras que el Markup está disponible para cualquier revista que no pertenezca a la red SciELO. Lo importante de señalar en este aspecto, es si los editores son los usuarios directos de estas herramientas, en el caso del Markup, es mayormente utilizado por empresas que realizan la marcación y que son certificadas por los diferentes sitios SciELO de cada país, las cuales son contratadas por las revistas para que hagan el trabajo de marcación, es decir, es una práctica que incentiva la tercerización de un trabajo que debería ser capaz de hacer el propio editor de la revista si así lo desea y no verse obligado a contratar a un tercero, por la complejidad de la herramienta que tiene a su disposición. En el caso de Marcalyc, es una herramienta de marcación que puede ser usada por un editor que no tenga conocimientos sobre formatos de publicación electrónica, ya que Redalyc ha puesto a disposición de los editores un sistema completo de capacitación sobre el uso de dicha herramienta (<http://marcalyc.redalyc.org/ayuda/>), es decisión del editor si realiza el trabajo de marcación o desea tercerizar dicho proceso.

Finalmente, se debe destacar que Marcalyc incluye la opción de generar el formato XML-JATS bajo los lineamientos del esquema de publicación de SciELO (SPS v. 1.6), con lo cual sería necesario que el editor haga un solo trabajo de marcación para cumplir con lo solicitado por ambos metaeditores.

Bibliografía

- Aguado-López, Eduardo, & Becerril-García, Arianna (2017). Redalyc 15 Años. Fin de un ciclo, inicio de otro. Recuperado de: <https://youtu.be/SWex1MGmwzQ>.
- Aguado-López, Eduardo; Becerril-García, Arianna; Guzmán-Useche, Eliana; Segundo-Escobar, Alma, & Rodríguez-Contreras, Fernando. (2016). Manual de Usuario Marcalyc. 2ª. Ed. Toluca, México. Sistema de Información Científica Redalyc-UAEM.
- Babini, Dominique, y Machin-Mastromatteo, Juan D. (2015). Latin American Science is meant to be open access: Initiatives and current challenges. Information Development, 31(5), 477-481. Recuperado de: <https://doi.org/10.1177/0266666915601420>.
- Crow, Ryam. (2013), Sustainability of Open Access Services. Phase 3: The Collective Provision of Access Resources, Washington, USA: SPARC.
- Herrera, Luis A.; Ferracutti, Víctor M.; Martínez, Fernando A., & Mitzig, Natalia 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>.
- Guzmán-Useche, Eliana, & Rodríguez-Contreras, Fernando. (2016). Sustentabilidad de las iniciativas latinoamericanas de publicación de revistas científicas en acceso abierto utilizando el estándar XML JATS: el caso de SciELO. Biblios, (64), 15-32, Recuperado de: <https://doi.org/10.5195/biblios.2016.290>.
- 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/>.
- Olvera-García, Jorge; Gentili, Pablo; Aguado-López, Eduardo; Babini, Dominique, & Becerril-García, Arianna. (2015). Catálogo de Revistas de Ciencias Sociales y Humanidades Redalyc-Clacso. 1a ed. Ciudad Autónoma de Buenos Aires: CLACSO; Toluca, México: Redalyc-UAEM.
- Rozemblum, Cecilia, Bava, Leslie, Unzurrunzaga, Carolina, y Banzato, Guillermo (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>.
- SCImago. (2007). SJR - SCImago Journal & Country Rank. Recuperado de: <http://www.scimagojr.com>.
- Revista Digital Universitaria, "El papel de Redalyc en el Acceso Abierto. Entrevista con Eduardo Aguado López" Revista Digital Universitaria [en línea]. 1 de octubre de 2014, Vol. 15, No.10 [Consultada:]. Recuperado de: <http://www.revista.unam.mx/vol.15/num10/art78/index.html>. ISSN: 1607-6079.

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 Científicas Venezolanas REVENCYT. Universidad de Los Andes, Mérida, Venezuela.

frodriguezcv@gmail.com

Received - Recibido: 2018-03-13

Accepted - Aceptado: 2019-07-24



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](#).