

Propuesta de herramienta informática para la conversión a formatos de publicación en la gestión editorial universitaria

Computer tool proposal for publishing formats conversion in university editorial management

Alvarez Carrera, Reinaldo; Bermúdez Torres, Iseidy; Alemán Assef, Lourdes; Ulloa Paz, Emeria

Reinaldo Alvarez Carrera reinaldoac@unica.cu
Universidad de Ciego de Ávila Máximo Gómez Bález, Cuba

Iseidy Bermúdez Torres iseidy@unica.cu
Universidad de Ciego de Ávila Máximo Gómez Bález, Cuba

Lourdes Alemán Assef lourdesea@unica.cu
Universidad de Ciego de Ávila Máximo Gómez Bález, Cuba

Emeria Ulloa Paz emeriam@unica.cu
Universidad de Ciego de Ávila Máximo Gómez Bález, Cuba

Estrategia y Gestión Universitaria

Universidad de Ciego de Ávila, Cuba

ISSN-e: 2309-8333

Periodicidad: Semestral

vol. 10, núm. 1, 2022

dailenn91@gmail.com

Recepción: 20 Enero 2022

Aprobación: 28 Febrero 2022

URL: <http://portal.amelica.org/ameli/journal/780/7803924013/>

Resumen: El presente estudio se centró en la búsqueda de mejoras para el proceso de conversión de documentos para publicación en la gestión editorial de revistas electrónicas, pertenecientes a la Universidad de Ciego de Ávila Máximo Gómez Bález. Se realizó un análisis de las tecnologías más usadas en la conversión a formatos de publicación, y se llevó a cabo una evaluación comparativa entre éstas y la tecnología Pandoc, para convertir a los formatos PDF y HTML, donde se incluyeron dos formatos candidatos a incorporar en futuras publicaciones (EPUB y XML JATS). Como resultado, se determinó la factibilidad de construir una aplicación informática que presentara una interfaz de fácil uso, con opciones de configuración y gran libertad de personalización para los requerimientos de cada revista, mediante la utilización de archivos externos de plantillas y metadatos. Por último, se presentó y explicó un prototipo de diseño de la interfaz de la aplicación de conversión de documentos usando la tecnología Pandoc.

Palabras clave: gestión editorial, conversión de documentos, evaluación de herramientas, pandoc, aplicación informática.

Abstract: The present study was centered in searching for improvements for documents for publication conversion process in editorial management of electronic journals, in the Máximo Gómez Bález University of Ciego de Ávila. An analysis of the most used conversion technologies for publishing formats was carried out, and a comparative evaluation was performed between these and the Pandoc technology for converting to PDF and HTML formats, plus two candidate formats to be incorporated in future publications (EPUB and XML JATS). As a result, the feasibility of building a computer application that presented an easy-to-use interface, with configuration options and great freedom of customization for the requirements of each journal, through the use of external templates and metadata files, was determined. Finally, a design prototype of the document conversion application interface using Pandoc technology was presented and explained.

Keywords: editorial management, document conversion, pandoc, tools evaluation, computer application.

Introducción

La gestión editorial de las revistas electrónicas universitarias es un proceso complejo, que incluye varios aspectos: gestión de publicaciones inéditas, revisión por pares, corrección de los originales, servicios de diseño e indexación en bases de datos, entre otros. El perfeccionamiento de cada una de las aristas que componen este proceso, contribuye a mejorar la visibilidad de los resultados científicos de las universidades cubanas y apoya la labor docente.

En este sentido, la proyección y cumplimiento de los parámetros de calidad exigidos por las bases de datos reconocidas, es una de las tareas fundamentales de la gestión editorial universitaria. Actualmente, el ingreso de las revistas universitarias cubanas en bases de datos reconocidas es una de las tareas fundamentales. Uno de los requisitos que se exigen para lograr la indexación es lograr publicaciones en diversos formatos (PDF, HTML, EPUB, XML JATS, JATS4R, TEI XML). Por esta razón, es notable la fase de conversión del documento hacia todos los formatos de publicación. Este proceso puede presentar grandes desafíos, pues se requieren documentos finales que cumplan con las normas de la revista en cuestión.

Existen muchas revistas electrónicas que han extendido la variedad de los formatos en que se realizan sus publicaciones. Esto ayuda al lector pues le brinda la posibilidad de visualizar las publicaciones en variedad de dispositivos de lectura. Por otra parte, también contribuye al crecimiento y visibilidad de la revista, ya que algunos formatos son ampliamente utilizados para intercambiar información entre revistas, bibliotecas y para mejorar la indexación de las publicaciones en buscadores electrónicos.

Actualmente, existe una gran inclinación hacia el uso de la marcación en la publicación de documentos digitales. Una de las mayores ventajas de la marcación radica en incorporar una estructura semántica en manuscritos, siendo el formato de archivo XML uno de los más usados.

Un reporte de proyecto en (University of Edinburgh, 2019), menciona que el uso de XML en publicaciones de libros y revistas brinda las siguientes ventajas:

- Contenido académico legible a las computadoras e independiente del diseño
- Formato más flexible y reusable para una variedad de formatos (PDF, HTML, EPUB)
- Capacidad de búsqueda mejorada
- Accesibilidad
- Preservación
- Uso de la minería de textos
- Enriquecimiento de contenido a través de multimedia
- Etiquetado semántico

Dicho reporte se basa en el uso de una nueva estructura XML llamada XML JATS (*Journal Article Tag Suite*)¹, la cual fue diseñada con el propósito de marcar documentos científicos, en la dirección de aprovechar estas ventajas.

Adicionalmente, una conferencia en Library Publishing Forum 2020 (Bezsheiko et al., 2020), expone que los autores buscaban lograr una marcación JATS XML que contuviera: metadatos, texto completo y referencias. Uno de sus propósitos fundamentales fue identificar, de forma precisa, los elementos que componen el texto para contribuir a la lectura del documento, pero también para su indexación.

Este trabajo tiene como objetivos los siguientes: identificar las herramientas de conversión más usadas en las revistas electrónicas de la Universidad de Ciego de Ávila Máximo Gómez Báez; realizar un análisis comparativo de resultados de conversión de la herramienta Pandoc y de las herramientas de conversión más usadas, donde se evalúen ventajas y desventajas de las mismas en cuanto a la conversión desde el formato de Word DOCX hacia los siguientes formatos (PDF, HTML, XML JATS, EPUB); y proponer una solución para mejorar la gestión del proceso de conversión a estos formatos. Ella consiste en la creación de una aplicación informática para Windows, que permita usar la potencialidad de la herramienta Pandoc (Macfarlane & Pandoc Collaborators, 2022) a través una interfaz intuitiva y sencilla, teniendo en cuenta un diseño que permita lidiar con cambios venideros en la tecnología subyacente (Pandoc).

Desarrollo

La primera etapa se compuso de un análisis de los formatos que habían sido usados en las publicaciones anteriores de las revistas: *Educación y Sociedad*; *Estrategia y Gestión Universitaria*; *Universidad & Ciencia*. Todas han recibido los artículos en los formatos DOC y DOCX. Los formatos que usaron para publicar fueron: PDF como formato primario y HTML y XML como formatos secundarios. No obstante, el XML usado no se corresponde con el estándar JATS, y por esa razón no lo suficientemente adecuado para los procedimientos de indexación y minería de datos en textos. Entonces, se procedió a evaluar las herramientas más usadas para convertir a estos formatos.

Aplicaciones con licencia de uso pagada

Por parte de las aplicaciones de pago, destaca el Microsoft Word en sus versiones desde 2010 hasta 2019. Una evaluación en *PCMag* (Mendelson, 2022) muestra la suite de Microsoft Office como la mejor entre las más destacadas, con una puntuación de 4.5 y solamente igualada por Apple iWork.

Aplicaciones con licencia de uso gratuita

Entre las aplicaciones gratis para la creación, edición y conversión de documentos DOC y DOCX, destacan LibreOffice y OpenOffice. Valoraciones comparativas de versiones recientes de ambas aplicaciones (Baxter, 2021; George, 2021; Velasco, 2021) concluyen que LibreOffice ofrece mejores funcionalidades y soporte. Además, una pizarra comparativa confeccionada por *Finances Online* (Finances Online, 2021), también muestra puntuaciones superiores para LibreOffice respecto a OpenOffice.

Evaluación de las herramientas seleccionadas

Una vez revisada la documentación de las nuevas características de Microsoft Word 2021 (Microsoft Team, 2022), se determinó que no eran significativos los cambios concernientes a los formatos de archivo respecto a las dos versiones anteriores. En este sentido, se escogieron dos de las versiones más actuales de Microsoft Word (2016 y 2019) y la más reciente de LibreOffice (7.2.5) hasta mediados de febrero de 2022. Luego, se procedió a estudiar el nivel de soporte de Microsoft Word y LibreOffice en cuanto las siguientes funcionalidades:

- Soporte para los formatos requeridos para publicación
- Soporte para añadir estructura de marcación (bookmarks) al documento resultante
- Soporte para inclusión de metadatos en documento resultante
- Soporte para fórmulas en texto
- Soporte para imágenes autocontenidas

Los documentos de origen se encontraban en formato DOCX, y tenían las siguientes características:

1. No contenían objetos de título de tabla
2. No contenían objetos de título de imagen
3. No contenían hipervínculos a objetos dentro del documento
4. No contenían referencias bibliográficas con archivos de referencias externos
5. No contenían estilos personalizados para los estilos de encabezado

Tabla 1

Características soportadas por Microsoft Word 2016, 2019 en la conversión del formato DOCX hacia los formatos PDF, HTML, XML JATS y EPUB

	PDF	HTML	XMLJATS	EPUB
Soporta conversión	x	x	-	-
Soporta marcación (bookmarks)	x	toc	-	-
Soporta inclusión de Metadatos	x	-	-	-
Fórmulas como texto	x	-	-	-
Imágenes autocontenidas	x	-[2]	-	-

Elaborado por los autores

Leyenda: x (soporta la característica), - (no soporta la característica), toc (soporta la característica en forma de tabla de contenidos)

Tabla 2

Características soportadas por Libre Office 725 en la conversión desde los formatos DOCX y ODT hacia los formatos PDF HTML XML JATS y EPUB

	PDF	HTML	XML JATS	EPUB
Soporta conversión	x	x[3]	-	p
Soporta marcación (bookmarks)	x	toc	-	x
Soporta inclusión de Metadatos	x	x	-	p
Fórmulas como texto	x	d	-	-
Imágenes autocontenidas	x	x	-	x

Elaborado por los autores

Leyenda: x (soporta la característica), - (no soporta la característica), toc (soporta la característica en forma de tabla de contenidos), p (la conversión tiene una cobertura parcial de la estructura del DOCX original), d (la conversión es deficiente)

Los resultados de la comparación, mostraron de forma general, que el formato PDF fue el de mejor soporte, seguido por el HTML. El soporte fue nulo para el caso del XML JATS, mientras que el EPUB tuvo un soporte aceptable en LibreOffice, aunque no detectó correctamente las viñetas y tampoco identificó las fórmulas matemáticas.

Resultados de conversión sobre Open Journal Systems (OJS)

Se comprobó la conversión del formato DOCX al formato XML JATS, mediante el uso de herramientas integrables en OJS. Las pruebas realizadas fueron efectuadas sobre la versión 3.3.0-8 de OJS, con el plugin docxConverter (Bezsheiko, 2021) en su versión 1.1.1 para la conversión a XML JATS y el plugin texture (Whitanage, 2021) en su versión 2.4.3-8 para la edición del documento en formato XML JATS.

Los resultados de conversión son aceptables, pero se detectaron problemas:

- En ocasiones insertó espacios en blanco
- En ocasiones no detectó las imágenes y hubo que recargarlas
- En ocasiones no guardó los cambios realizados

Es necesario resaltar que el Plugin textura, para edición de XML JATS sobre OJS, puede producir un acabado bastante bueno si se realiza una edición minuciosa. Aun así, presenta las siguientes limitaciones:

- No permite aplicar alineación al texto
- Las referencias hay que definir las de forma manual o importarlas usando el un formato CLS JSON. No obstante, sólo muestra las referencias en el estilo bibliográfico Vancouver y en la experiencia de trabajo durante este estudio fue necesario las referencias que se importadas tuvieran especificado un valor en el campo DOI
- Este plugin está terminando su ciclo de desarrollo, pues como se planteó en (Bezsheiko et al., 2020), la comunidad se está moviendo ahora a Prosemirror (Prosemirror Team and Collaborators, 2022)

Resultado de conversión con Pandoc

En este punto, se comenzó la preparación para realizar conversiones mediante la herramienta Pandoc en su versión 2.17.1.1. Esta cuenta con una serie de utilidades de línea de comandos que permiten realizar conversiones bastante completas desde 39 formatos origen (más uno personalizado.) hacia 56 formatos destino (más uno personalizado.).

Esta herramienta puede ser enriquecida en funcionalidades de conversión mediante el uso de filtros que el usuario puede crear en determinados lenguajes de programación permitidos por esta tecnología (Macfarlane, 2022). Además, los formatos de salida son altamente configurables mediante el uso de plantillas personalizadas para construir el documento final en su formato correspondiente. También, se pueden usar archivos de configuración en formato YAML y XML según sea el formato de destino y de forma particular, es posible personalizar el acabado del documento haciendo uso los siguientes recursos:

- Formato PDF: instrucciones LaTeX (Mittelbach et al., 2004) y scripts KOMA (Kohm, 2021) en archivos YAML
- Formato HTML: hojas de estilo en cascada (CSS), instrucciones en código HTML (se refiere a un formato de origen y un formato de destino que tiene una estructura personalizada por el usuario)
- Formato EPUB: hojas de estilo en cascada (CSS), instrucciones en código HTML

Estos recursos pueden ser usados directamente en el cuerpo de un archivo en formato Markdown de Pandoc⁴, o como parte de archivos de configuración en formato YAML. Adicionalmente, se puede incluir la información correspondiente a los metadatos en un archivo YAML para varios formatos destino, entre ellos: PDF, HTML, EPUB y XML JATS. De forma alternativa se pueden especificar los metadatos en forma de parámetros pasados al comando que ejecuta la conversión.

Para lograr que los documentos tuvieran el acabado más cercano a las normas de las revistas, fue necesario convertir el formato DOCX a formato Markdown de Pandoc. Luego, se convirtió desde este formato hacia el resto y se obtuvieron conversiones automáticas altamente satisfactorias. A continuación, se describe en detalle cada formato con las características positivas y negativas que se lograron identificar.

Tabla 3
Evaluación de los resultados de conversión con Pandoc

Formato	Positivo	Negativo
HTML	<ol style="list-style-type: none"> 1. Las fórmulas se muestran en forma de texto y con total fidelidad. 2. Las imágenes pueden ser incrustadas en el cuerpo del HTML. 3. Si se usa MathJax o KaTeX para generar las fórmulas, estas pueden visualizarse en la mayoría de los navegadores. 	<ol style="list-style-type: none"> 1. Algunos párrafos no se alinearon a la derecha (corregible con código HTML en el archivo Markdown). 2. El espaciado entre títulos de sección es muy grande (corregible con código HTML en el archivo Markdown). 3. Algunos párrafos no tomaron la sangría (corregible con código HTML en el archivo Markdown). 4. Algunos objetos SmartArt de Word no son detectados correctamente (corregible si antes se convierte a imagen). 5. La generación de fórmulas usando KaTeX incrementa el tamaño de archivo en 4 veces aproximadamente.
EPUB	<ol style="list-style-type: none"> 1. Las viñetas son detectadas fielmente 2. Las fórmulas son capturadas fielmente 	<ol style="list-style-type: none"> 1. Algunos párrafos no se alinearon a la derecha (corregible). 2. El espaciado entre títulos de secciones muy grande (corregible). 3. Algunos párrafos no tomaron la sangría (corregible).
XML JATS	Permite incluir gran cantidad de metadatos	No incrusta la información de las imágenes, sino que las toma como referencias a archivos externos, excepto cuando se convierte desde un archivo origen en formato HTML con imágenes autocontenidas.

Elaborado por los autores

Es importante destacar que el formato XML JATS no incluye estilos, pues no fue diseñado con ese propósito, sino para intercambiar información entre revistas y describir el contenido gráfico y textual de documentos (National Information Standards Organization, 2019).

Resultado de conversión con LibreOffice

La tabla a continuación muestra los resultados de conversión mediante el programa LibreOffice. Es de resaltar, que no fue posible detectar correcciones para las dificultades detectadas en la conversión desde LibreOffice, pero sí se hallaron correcciones aplicables para la mayoría de las dificultades detectadas al convertir desde Pandoc

Tabla 4
Evaluación de los resultados de conversión con LibreOffice

Formato	Positivo	Negativo
PDF	El acabado tiene una gran fidelidad al documento original	
HTML	Convierte bien los objetos de gráfico	1. No detecta la alineación de columnas a la derecha 2. La conversión de fórmulas es deficiente
EPUB	Detecta bien la alineación del texto en las tablas	1. No detecta las viñetas (numeradas y no numeradas) 2. No detecta las fórmulas de Office 3. No detecta los SmartArt de Word 4. No permite incrustar imágenes en el cuerpo del archivo
XML JATS		No es posible obtener este formato

Elaborado por los autores

Inclusión de metadatos en el archivo DOCX u ODT

Word 2016, 2019 soporta la inclusión de metadatos personalizados en el archivo DOCX. El menú *Archivo/Propiedades/Propiedades avanzadas* permite especificar las propiedades más comunes y en su pestaña *Resumen* y se pueden agregar metadatos personalizados en la pestaña *Personalizar*.

LibreOffice 7.2.5 soporta la inclusión de metadatos personalizados en el archivo DOCX y ODT. El menú *Archivo/Propiedades* permite especificar las propiedades más comunes y en su pestaña *Descripción* y se pueden agregar metadatos personalizados en la pestaña *Propiedades personalizadas*.

Pandoc permite la inclusión de metadatos usados comúnmente definidos y también permite definir metadatos personalizados para los formatos DOCX, ODT y PPTX. Adicionalmente, hasta la realización de este estado, la versión más actual de Pandoc proveía más de 50 etiquetas^[5] definidas describir el formato XML JATS.

Diseño propuesto para la herramienta de conversión

Se propone una interfaz visual con estructura de Formulario, que se distinga por las siguientes ventajas:

1. Permitir al usuario acceder a los recursos de conversión de Pandoc sin la necesidad de aprender la sintaxis de línea de comandos
2. Presentar una interfaz intuitiva y simplificada, pero que permita un nivel considerable de personalización
3. Agrupar por secciones los datos que entra el usuario, teniendo en cuenta el propósito de los mismos en el acabado de los documentos resultantes de la conversión

4. Permitir el mayor nivel de personalización posible desde archivos de configuración externos, para que sea posible lidiar con futuras actualizaciones tecnológicas de Pandoc
5. Permitir la posibilidad de convertir a varios formatos a la vez, especificando sólo una vez los datos de cada artículo

La desventaja principal, aunque de poco impacto, es que el usuario debe aprender una sintaxis mínima de Markdown de Pandoc, LaTeX y HTML para lograr un acabado bien ajustado a sus requerimientos. Esta desventaja se debe a que el lenguaje Markdown de Pandoc aún está en desarrollo y no es capaz de capturar ciertas características avanzadas de uso común en otros formatos como DOCX y ODT, aunque es posible que esta brecha se vaya cerrando a medida que la tecnología es perfeccionada.

Para que la herramienta funcione correctamente, se requiere preparar el entorno de trabajo de la forma siguiente:

1. La herramienta Pandoc (posee un asistente de instalación sencillo)
2. La herramienta ProTEX que contiene la instalación de LaTeX, con la posibilidad de agregar todos los paquetes necesarios para la conversión a PDF (de forma online u offline). Se recomienda la versión 3.1.10
3. Editor de archivos Markdown (portable o instalable). Puede ser un Bloc de notas, pero se recomienda Visual Studio Code con la extensión Pandoc Markdown desarrollada por el usuario rsolano o la herramienta PanWriter
4. Visor de EPUB para evaluar el acabado de los documentos EPUB. Se recomienda Calibre en su versión más reciente
5. La carpeta del documento origen debe contener determinados recursos

Recursos que debe contener la carpeta del documento origen:

- Plantilla latex para generar el documento PDF
- Archivo de configuración pdf.yaml para la creación del PDF
- Archivo de pdf_encabezado.yaml que contiene la estructura en código LaTeX para construir el encabezado del PDF
- Archivo de configuración html.yaml para la creación del HTML
- Archivo de estilo plantilla.css para la aplicación de estilos visuales al documento HTML
- Archivo encabezado.html con la estructura en código HTML para construir encabezado del HTML o del EPUB
- Archivo de configuración epub.yaml para la creación del EPUB
- Archivo de estilo epub_plantilla.css para la aplicación de estilos visuales al documento EPUB
- Archivo de configuración xmljats.yaml para la creación del XML JATS
- Archivo xmljats_tags.txt con la estructura de las etiquetas admitidas por Pandoc
- Carpeta katex con los archivos necesarios para generar las fórmulas usando KaTeX

- Carpeta mathjax con archivos necesarios para generar las fórmulas usando MathJax

A continuación, se muestran imágenes de un prototipo inicial de la herramienta de conversión:

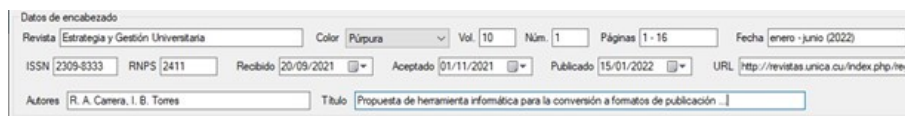


Figura 1

Sección de encabezado

La sección de encabezado recoge los datos necesarios de los PDF y HTML (se pueden incluir opcionalmente en el formato EPUB)
Elaborado por los autores

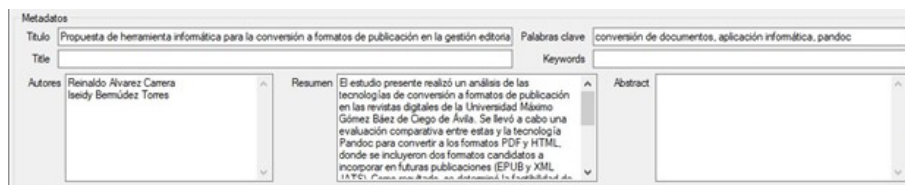


Figura 2

Sección de metadatos

La sección de metadatos, recoge los principales datos a incluir en los archivos PDF, HTML, EPUB y XML JATS
Elaborado por los autores

Figura 3. Sección documento de origen.

La sección documento de origen, abre un diálogo de apertura de archivo, que permite seleccionar. Para el prototipo inicial se propone que sean en formato Markdown y HTML.

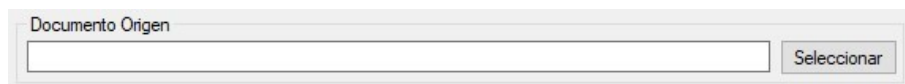


Figura 3

Sección documento de origen

La sección documento de origen, abre un diálogo de apertura de archivo, que permite seleccionar. Para el prototipo inicial se propone que sean en formato Markdown y HTML
Elaborado por los autores

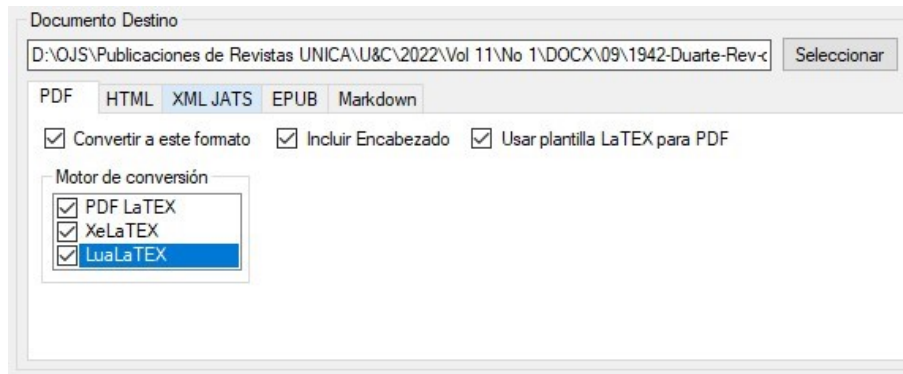


Figura 4
Sección de documento de destino

La sección de documento de destino abre un diálogo de guardado que permite seleccionar una ubicación, un formato y un nombre para el archivo destino. Además, contiene una pestaña para las opciones de cada formato, entre las que se encuentran:

- Convertir a este formato donde se especifica si se realizará la conversión a ese formato y abre la posibilidad de personalizar la selección de formatos destino
- Incluir encabezado (PDF, HTML, EPUB)
- Usar plantilla de configuración (PDF, pero puede extenderse a los demás formatos)
- Selección de motores de conversión (PDF). En caso de seleccionar más de uno, se generarán archivos con el nombre especificado en el diálogo de guardado con la adición de la inicial del nombre de motor de conversión (Ejemplo: p para PDF LaTeX)
- Generar HTML autocontenido (HTML)
- Permitir párrafos vacíos (HTML)
- Usar hoja de estilos (HTML, EPUB)
- Editor en forma de árbol que permita la especificación de las etiquetas permitidas de XML JATS. Este editor cargará la estructura desde un archivo texto que contenga la estructura etiquetas con una etiqueta por línea y asignando grupos de espacios para cada nivel (Figura 5)
- Los archivos de estilo YAML y HTML contendrán etiquetas encerradas entre símbolos de \$ (Ejemplo \$pRevista\$), las cuales serán sustituidas por sus valores correspondientes antes de generar el documento final y así mostrar los valores correctos

Elaborado por los autores

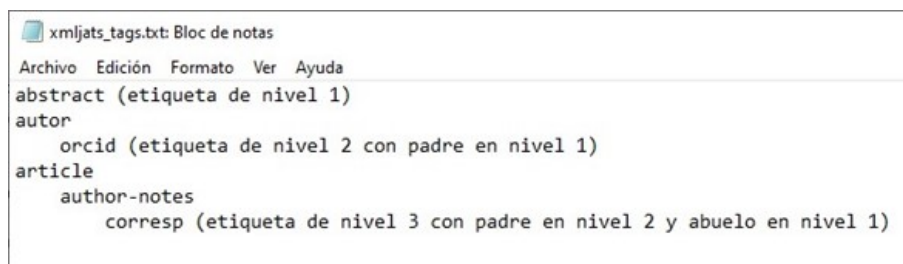


Figura 5
Estructura propuesta para el archivo que contiene la configuración de las etiquetas para XML JATS

El texto entre paréntesis no se escribirá en el archivo, se ha colocado con el objetivo de comprender el nivel que corresponde a cada etiqueta.

Elaborado por los autores

```
Convirtiendo archivo 1 de 1 ...
pandoc.exe "1942-Duarte-Rev-cs.md" 4 pdf -pdf-engine=pdflatex -metadata-file=pdf_meta.yaml -template=plantilla.tex -o "1942-Duarte-Rev-cs-ultimop.pdf"
Código de error: 0
```

Figura 6

Sección de logs de ejecución

Por último, la sección de logs de ejecución que muestra información del progreso de todo el proceso de conversión, con el objetivo de monitorearlo para detectar errores con sus respectivos mensajes e intentar corregirlos

Elaborado por los autores

Conclusiones

El uso de la herramienta Pandoc, mostró ser una solución factible al problema de conversión de documentos para publicación en las revistas de la Universidad de Ciego de Ávila. Esta tecnología, permite obtener documentos con la calidad suficiente para sustituir herramientas de conversión como Microsoft Word y LibreOffice.

La aplicación informática propuesta, brindará una interfaz que permitirá hacer uso de las funcionalidades de Pandoc requeridas para obtener los resultados deseados, ayudando al usuario (editor) a ganar en simplicidad, agilidad y precisión durante el proceso de conversión.

Disponer de esta nueva herramienta contribuye a lograr la indexación de las revistas en nuevas bases de datos, ya que permite la conversión de las publicaciones a diversos formatos, y es este uno de los requisitos de calidad exigidos a nivel internacional.

Referencias bibliográficas

- Baxter, D. (2021). *Libreoffice vs OpenOffice*. TechRadar The Source for Tech Buying Advice. Recuperado de <https://www.techradar.com/news/libreoffice-vs-openoffice>
- Bezsheiko, V. (2021). docxConverter plugin for DOCX to XML format conversion in OJS 3 (1.1.1). Recuperado de <https://github.com/Vitaliy-1/docxConverter>
- Bezsheiko, V., Withanage, D., & MacGregor, J. (2020). XML Publishing Workflows. Tools, Practices and the Future. Library Publishing Forum 2020. Recuperado de [https://librarypublishing.org/wp-content/uploads/2020/05/ PKP-XML-Publishing-Workflows.pdf](https://librarypublishing.org/wp-content/uploads/2020/05/PKP-XML-Publishing-Workflows.pdf)
- Finances Online. (2021). Compare Apache OpenOffice vs LibreOffice. FinancesOnline Reviews for Bussines. Recuperado de [https:// comparisons.financesonline.com/apache-openoffice-vslibreofficeç](https://comparisons.financesonline.com/apache-openoffice-vslibreofficeç)
- George, A. (2021). LibreOffice vs. OpenOffice: Which one is right for you? DigitalTrends. Recuperado de <https://www.digitaltrends.com/computing/libreoffice-vs-openoffice/>
- Kohm, M. (2021). *KOMA-Script 3.28 The Guide* (2021st-11th-09 ed.).
- Macfarlane, J. (2022). *Pandoc User 's Guide* (January 30). Recuperado de <https://pandoc.org/MA-NUAL.pdf>
- Macfarlane, J., & Pandoc Collaborators. (2022). Pandoc. A universal document converter (2.17.1.1). Recuperado de <https://pandoc.org>

- Mendelson, E. (2022). The Best Office Suites for 2022. PCMag. Recuperado de <https://www.pcmag.com/picks/the-best-office-suites>
- Microsoft Team. (2022). *What's new in Word 2021 for Windows*. Microsoft. Recuperado de <https://support.microsoft.com/en-us/office/what-s-new-in-word-2021-for-windows-8f0b9117-8cf4-4677-8cd5-d848dd0e1ac9>
- Mittelbach, F., Goossens, M., Braams, J., Carlisle, D., & Rowley, C. (2004). *The LaTeX Companion Second Edition* (J. Andre, B. Beeton, D. Brailsford, T. Bray, P. Flynn, L. Lamport, C. Rowley, R. Rubinstein, & P. Stiff (eds.); 2nd ed.). Addison-Wesley.
- National Information Standards Organization. (2019). *ANSI/NISO Z39.96-2019, JATS: Journal Article Tag Suite, version 1.2*. Recuperado de <http://www.niso.org/publications/z3996-2019-jats>
- Prosemirror Team and Collaborators. (2022). *ProseMirror. A toolkit for building rich-text editors on the web*. Recuperado de <https://prosemirror.net/>
- University of Edinburgh. (2019). *Support for XML-based publishing in OJS*. Recuperado de https://www.ed.ac.uk/files/atoms/files/xml_publishing_in_ojs_-_project_summary_user_guide_0.pdf
- Velasco, R. (2021). *Utilidades OpenOffice vs LibreOffice: ¿Cuál es la mejor alternativa a Office?* SoftZone. Recuperado de <https://www.softzone.es/programas/utilidades/diferencias-openoffice/libreoffice/>
- Whitanage, D. (2021). *texture plugin for XML JATS edition on OJS (2.4.3-8)*. Recuperado de <https://github.com/pkp/texture>

Notas

- 1 Es una aplicación del estándar NISO Z39.96-2019 que provee un formato XML para que los editores y archivos puedan intercambiar contenido de las revistas. Este estándar define un conjunto de elementos y atributos que pueden ser usados describir contenido gráfico y textual de artículos de revista y de otros materiales (National Information Standards Organization, 2019)
- 2 De forma nativa Word convierte a formato HTML autocontenido pero con estructura y extensión mht. Este formato no es admitido por OJS como formato de publicación pues requiere que los archivos tengan estructura y extensión html o htm
- 3 Las viñetas representadas con figuras no se convierten apropiadamente
- 4 Es el formato base de Pandoc, a partir del cual se pueden lograr los mejores resultados cuando se convierte hacia otros formatos. Este formato constituye una modificación del Markdown original, al cual se agregaron nuevas funcionalidades propias de Pandoc
- 5 La lista de etiquetas más actualizada se encuentra en el enlace <https://pandoc.org/jats.html>