



EXPORTACIÓN DE COLECCIONES DOCUMENTALES PARA EL SISTEMA ARCHIVENHIS.



Autora: Lisbet Milagros Guerra Cantero

Universidad de las Ciencias Informáticas,
Carretera a San Antonio de los Baños, km 2 ½, Boyeros, La Habana, Cuba.

Autor para la correspondencia: lmcantero@grm.uci.cu

RESUMEN

El Sistema de Gestión de Documentos de Archivos para el Archivo General de la Nación de la República Bolivariana de Venezuela, nombrado ArchiVenHIS, se concibió para ser utilizado solamente desde las instalaciones de la institución, para apoyarla en su misión de preservar y difundir la documentación que representa la memoria histórica venezolana, sin embargo los investigadores que comenzaron a hacer uso del sistema identificaron que no les era posible extraer las descripciones de documentos o un conjunto de ellas con sus representaciones digitales. En la presente investigación se desarrolló un módulo para el ArchiVenHIS que permite crear colecciones documentales en las cuales se guardan las descripciones de los documentos de archivos con sus representaciones digitales asociadas, las que pueden ser exportadas en un paquete compactado para que puedan ser consultadas fuera de la institución. Se realizó el análisis y diseño del módulo utilizando la metodología de desarrollo de software RUP, generando los artefactos correspondientes para documentar el proceso de desarrollo y el diseño y aplicación de las pruebas de caja negra, donde se constató que el módulo cumple con las especificaciones requeridas.

Palabras clave: colección documental, descripción de documento, EAD, ISAD (G), XML

ÍNDICE

INTRODUCCIÓN.....	4
MATERIALES Y MÉTODOS	4
RESULTADOS Y DISCUSIÓN	10
APORTE Y NOVEDAD.....	13
CONCLUSIONES.....	13
REFERENCIAS.....	14

INTRODUCCIÓN

En el marco de la VI Reunión de la Comisión Mixta del Convenio Integral de Cooperación Cuba-Venezuela en el año 2006 se aprobó el Contrato de Cooperación “Uso y Aplicación de las TICs para el mejoramiento de la Gobernabilidad y Aumento de la Soberanía Tecnológica” entre la Universidad de las Ciencias Informáticas (UCI) y el Centro Nacional de Tecnologías de Información (CNTI). El objeto del Contrato era desarrollar sistemas de información basados en software libre que impulsaran a los organismos pertenecientes a la Administración Pública Nacional Venezolana en el proceso de migración a software libre. Entre los proyectos que se desarrollaron bajo el contrato se encuentra el Desarrollo e implementación de un Sistema de Gestión para el Archivo General de la Nación de la República Bolivariana de Venezuela (AGN), nombrado ArchiVenHIS. El sistema se concibió con la intención de ayudar al AGN con su misión de preservar y difundir la documentación que representa la memoria histórica venezolana, sin embargo los investigadores que comenzaron a hacer uso de él a partir del año 2007 identificaron la necesidad de extraer del sistema descripciones de documentos históricos para crear colecciones documentales¹ con la finalidad de organizar los documentos que pudieran ser relevantes para alguna investigación y poder hacer revisiones de ellos fuera de la institución. Se decide agregar un módulo que permita conformar y exportar colecciones de documentos de interés para los investigadores que puedan ser consultados en áreas externas al sistema.

Teniendo en cuenta lo expuesto anteriormente y la necesidad de dar solución a esta dificultad, surge en la presente investigación el siguiente problema científico: ¿Cómo lograr que los usuarios del Sistema ArchiVenHIS conformen colecciones documentales que puedan ser consultadas fuera del Archivo General de la Nación de la República Bolivariana de Venezuela?

De acuerdo con la problemática planteada se propone como objetivo general: Desarrollar un módulo para exportar del Sistema ArchiVenHIS descripciones de documentos de archivos para la conformación de colecciones documentales.

MATERIALES Y MÉTODOS.

Los **métodos teóricos** utilizados fueron el **Análisis histórico-lógico** y el **Analítico-sintético**.

El principal objetivo del método **Análisis histórico-lógico** en una investigación es constatar teóricamente cómo ha evolucionado un determinado fenómeno en un período de tiempo, en toda su trayectoria o en un fragmento temporal de la lógica de su desarrollo. En la presente investigación la utilización de este método se refleja en el estudio de la trayectoria histórica real de los sistemas de gestión de archivos, su condicionamiento y su desenvolvimiento y desarrollo en los diferentes períodos de la historia, así como en la gestión archivística en la investigación de su funcionamiento general en el proceso de las descripciones de los documentos históricos.

¹ La colección documental es la reunión artificial de documentos acumulados sobre la base de algunas características comunes, sin tener en cuenta la procedencia de dichos documentos. No debe confundirse con el fondo documental

El principal objetivo del método **Analítico-sintético** en una investigación es analizar las teorías, documentos, permitiendo la extracción de los elementos más importantes que se relacionan con el objeto de estudio. En la presente investigación se hace uso de este método basándose en la investigación, síntesis y extracción de los elementos más importantes relacionados con el proceso de descripciones de documentos históricos de los sistemas de gestión de archivos, además para sintetizar las normas y estándares que se utilizan en la descripción de los documentos, así como para la investigación y extracción de las características principales de los sistemas que se estudian, con el fin de comprender el objeto de estudio y la realización de la solución propuesta, para facilitar su estudio.

El **método empírico** seleccionado es la **Entrevista** su principal objetivo es una conversación planificada para obtener información, su uso constituye un medio para el conocimiento cualitativo de los fenómenos o sobre características personales del entrevistado. En la presente investigación se hace uso de este método en la realización de entrevistas a los especialistas del proyecto, dígase analistas, arquitectos y al jefe del proyecto, en aras de obtener información real referente al funcionamiento del negocio en el que se enmarca el problema.

El proyecto Sistema de Gestión de Documentos de Archivos ArchiVenHIS, surge en el año 2007 con el propósito de ofrecer al Archivo General de la Nación de la República Bolivariana de Venezuela (AGN) una herramienta desarrollada sobre tecnologías libres y estándares abiertos que automatiza el proceso de descripción archivística de la institución y la localización de los documentos con base en los metadatos descritos con el propósito de brindar servicios (préstamos, consulta digital, digitalización, fotocopias y transcripciones). Queda implantado como solución a la medida en el AGN en 2008. Su versión 1.0 se registró en el Centro Nacional de Derecho de Autor de Cuba (CENDA) en febrero 2010 con Número de Registro: 498-2010.

Sus principales funcionalidades:

- Facilita el proceso de descripción de los fondos documentales bajo la custodia de la institución según la norma ISAD (G).
- Permite asociar los documentos a sus representaciones digitales.
- Ofrece varias vías para la localización de los documentos tomando como base los metadatos incorporados durante la descripción (Búsqueda sencilla, Búsqueda avanzada, Explorar).
- Facilita la difusión de la memoria histórica bajo la custodia de la institución, pues permite la localización de documentos, así como el acceso a sus representaciones digitales por varios usuarios de forma concurrente, favoreciendo a su vez la conservación de los fondos documentales al evitar su manipulación directa.
- Permite obtener reportes acerca del proceso de descripción, de los documentos consultados, del estado de conservación de los fondos documentales, de los resultados de los tratamientos aplicados y de

los parámetros ambientales de los locales empleados para la custodia de los mismos, contribuyendo a la toma de decisiones en la institución.

La descripción de documentos es un medio utilizado por el archivero para obtener la información contenida en los documentos y ofrecerla a los interesados en ella. Persigue dos objetivos: dar información a los demás y facilitar el control al archivero.

La descripción de documentos es un proceso técnico archivístico que consiste en identificar y analizar los documentos con la finalidad de elaborar los auxiliares o instrumentos descriptivos (guías, inventarios, catálogos) que permitan conocer, localizar y controlar los fondos documentales de cada entidad. Estas descripciones permitirán a los usuarios e investigadores conocer cuáles están disponibles sin necesidad de acceder a los originales, lo que permitirá brindar además un eficiente y eficaz servicio de información.

La finalidad de la descripción archivística es identificar y explicar el contexto y el contenido de los documentos de archivo con el fin de hacerlos accesibles. **(Mena Mugica, 2005)**

Para lograr descripción archivísticas coherentes y explícitas se utiliza la Norma Internacional General de Descripción Archivística ISAD(G) por sus siglas en inglés, la cual se concibe como una guía general para la elaboración de descripciones archivísticas. Permite la creación de una representación precisa de los fondos y sus partes componentes mediante el proceso de captar, componer y organizar la información que sirve para identificar el material archivístico y explicar el contexto y sistemas archivísticos que lo produjeron.

La norma ISAD(G) define 26 elementos descriptivos que pueden combinarse entre sí para constituir la descripción de una unidad archivística, los cuales se agrupan en 7 áreas de información descriptiva. **(Corral, 1999)**

Área de Identificación: Contiene la información esencial para identificar la unidad de descripción.

- Código(s) de referencia.
- Título.
- Fechas(s) extremas de producción.
- Nivel de descripción.
- Extensión y soporte de la unidad de descripción (cantidad, volumen o tamaño).

Área de Contexto: Contiene la información relativa al origen y custodia de la unidad de descripción.

- Nombre(s) del productor(es).
- Historia Institucional/Reseña biográfica.
- Historia Archivística.
- Forma de ingreso.

Área de Contenido y Estructura: Contiene la información relativa al objeto y organización de la unidad de descripción.

- Alcance y contenido.
- Información sobre valoración, selección y eliminación.
- Nuevos ingresos.
- Sistema de organización.

Área de Acceso y Utilización: Contiene la información relativa a la accesibilidad de la unidad de descripción.

- Condiciones de acceso.
- Condiciones de reproducción.
- Lengua/escritura del material.
- Características físicas y requisitos técnicos.
- Instrumentos de descripción.

Área de Documentación Asociada: Contiene la información relativa a aquellos documentos que tienen una relación significativa con la unidad de descripción.

- Existencia y localización de los documentos originales.
- Existencia y localización de copias.
- Unidades de descripción relacionadas.
- Notas de publicaciones.

Área de Notas: Contiene información especial y aquella otra que no ha podido incluirse en ninguna de las demás áreas.

- Notas.

Área de Control de la Descripción: Contiene la información relativa al cómo, cuándo y quién ha elaborado la descripción archivística.

- Nota del archivo.
- Reglas o normas.
- Fechas(s) de la(s) descripción(s).

Para el intercambio de descripciones de documento de archivos se utiliza el Estándar de Descripción Archivística Codificada EAD por sus siglas en inglés, para la codificación y transferencia de descripciones en formato electrónico, con el propósito de intercambiar entre sistemas información de las descripciones de documentos electrónicos, el cual consiste en una Definición de Tipo de Documento (DTD), válida para el Lenguaje de Señalización General Normalizado (SGML), que dice qué estructura deben tener las descripciones de documentos de archivos, y las etiquetas que se pueden utilizar en su redacción.

El objetivo de EAD es normalizar el formato utilizado para crear, codificar e intercambiar descripciones de documentos de archivo en formato electrónico. **(Eito-Brun,2005)**

Se han establecido equivalencias entre los elementos definidos en ISAD (G) y los de EAD para garantizar que cualquier descripción EAD sea compatible con ISAD (G).

Tabla 1. Correspondencia entre ISAD (G) y EAD.

Elementos ISAD(G)	Descripción	Elemento EAD
Código de referencia.	Permite localizar físicamente la unidad de descripción en el fondo. Debe incluir el código del país codificado, seguido del código del archivo (indicado según las normas nacionales) y la signatura de la unidad documental propiamente dicha.	<repository> <physloc> <unitid>
Título.	Título descriptivo que identifica a la unidad de descripción.	<unittitle>
Fechas.	Fecha de creación del documento o de la unidad de descripción (fechas de constitución de la unidad de descripción o acumulación).	<unitdate>
Nivel de descripción.	Indica si se está describiendo una serie, expediente, fondo, subfondo o unidad documental.	<archdesc level=" ">
Extensión y soporte.	Número de unidades físicas o alternatively el espacio lineal o cúbico de la unidad de descripción.	<physdesc>
Productor.	Nombre de la institución, organización, familia o persona que produjo la unidad de descripción.	<origination>
Historia institucional/Reseña bibliográfica.	Datos biográficos o históricos sobre el productor. Permite contextualizar a la unidad de descripción.	<bioghist>
Historia archivística.	Información sobre los cambios de propiedad y custodia de la unidad de descripción, siempre que estos sean significativos.	<custodhist>
Fuente inmediata de adquisición/Forma de ingreso.	Procedencia, método y forma de ingreso de la unidad de descripción en el archivo. Se indicará la fuente inmediata de adquisición o transferencia.	<acqinfo>
Alcance y contenido.	Se refiere al contenido propiamente dicho de la unidad de descripción.	<scopecontent>
Valoración, selección y	Indicaciones sobre la valoración de los fondos, y las	<appraisal>

expurgo.	operaciones de selección y eliminación de documentos que se hayan realizado o que estén programadas.	
Nuevos ingresos.	Transferencias de documentos o ingresos previstos.	<accruals>
Organización.	Organización u ordenación de la unidad, sistema de clasificación.	<organization>
Condiciones de acceso.	Información sobre condiciones de acceso y utilización de los documentos desde los puntos de vista legal, físico o intelectual.	<accessrestrict>
Condiciones de reproducción.	Restricciones relativas a la reproducción de los documentos.	<userrestrict>
Lengua/escrituras del material.	Idiomas en los que se encuentra la documentación. También se utiliza este elemento para indicar tipos de letras o símbolos utilizados.	<langmaterial> langmaterial=" ">
Características físicas y requerimientos técnicos.	Indican si la unidad de descripción presenta características físicas que pueden afectar a la accesibilidad y utilización, así como la necesidad de utilizar alguna herramienta o dispositivo para leerla.	<phystech>
Instrumentos de descripción.	Se enumeran los instrumentos de descripción que ofrezcan información sobre la unidad de descripción: inventarios, índices, etc.	<otherfindaid>
Existencia y localización de los originales.	En aquellos casos en los que la unidad de descripción sea una copia, este elemento indicará dónde se encuentran los documentos originales.	<originalsloc>
Existencia y localización de copias.	Recogerá información sobre la existencia de copias de la documentación que se describe, cualquiera que sea su soporte físico.	<altformavail>
Unidades de descripción asociadas.	Establece relaciones con otras unidades de descripción.	<relatedmaterial> <separatedmaterial>
Nota de publicación.	Recoge referencias bibliográficas de publicaciones en las que se haya utilizado o analizado la unidad de descripción.	<bibliography>
Nota.	Cualquier información adicional sobre la unidad de descripción que no tenga cabida en el resto de elementos.	<odd>
Control de la descripción.	Nombre del archivero autor de la descripción, fecha de	Se corresponde con

	creación y normas en las que se basó.	elementos de la cabecera <eadheader>
--	---------------------------------------	---

RESULTADOS Y DISCUSIÓN

Para la realización del módulo de Exportación de colecciones documentales para el Sistema ArchiVenHIS se realizó un estudio del estado del arte de sistemas que implementaban colecciones documentales o el estándar EAD la exportación de descripciones de documentos realizados bajo la Norma ISAD (G) para analizar la forma en la cual estos sistemas implementan dichas funcionalidades.

Tabla 2. Tabla de patrones de comparación entre sistemas.

Patrones de comparación	ICA Atom	Archivo 3000	Archon	Albalá	ArchiVenHIS
Código abierto	X	-	X	X	X
Implementa norma ISAD (G)	X	X	X	X	X
Exporta descripciones utilizando EAD	X	X	X	-	-
Implementa colecciones documentales	-	-	X	-	-
Exportan colecciones con sus representaciones digitales asociadas	-	-	-	-	-

Luego de analizar los sistemas anteriormente mencionados se determinó que ICA AtoM y Archon son los que más se ajustan a la necesidades de la presente investigación, pues ambos sistemas permiten exportar descripciones archivísticas utilizando el estándar EAD y el sistema Archon en particular maneja el término de colección documental como el conjunto de documentos de archivos, sin embargo ninguno de estos sistemas permite exportar colecciones documentales que contengan las descripciones de los documentos de archivos con sus representaciones digitales. Se identificó además que no existe un estándar que permita exportar colecciones documentales con las descripciones archivísticas y sus representaciones digitales asociadas. Por tanto se decidió implementar funcionalidades que permitan conformar y exportar colecciones documentales que contengan las descripciones de los documentos de archivos, utilizando el estándar EAD con sus representaciones digitales en el sistema ArchiVenHIS.

Se identificaron como las principales funcionalidades para desarrollar la propuesta de solución:

Adicionar colección: permite adicionar una nueva colección documental del usuario autenticado.

Listar colección: permite listar el conjunto de colecciones documentales creadas por el usuario autenticado.

Modificar colección: permite modificar el nombre de una colección documental.

Eliminar colección: permite eliminar una colección documental.

Adicionar documento a colección: permite luego de realizar una búsqueda de documentos de archivos en el sistema, adicionarlos a una o varias colecciones documentales, pertenecientes al usuario autenticado.

Listar documento de colección: permite consultar los documentos que pertenecen a una colección documental específica.

Visualizar documento de colección: permite visualizar el archivo digital correspondiente al documento original.

Mover documento a colección: permite mover un documento de archivo hacia una colección específica.

Copiar documento a colección: permite hacer una copia de un documento de archivo en otra colección documental.

Eliminar documento de colección: permite eliminar un documento de archivo de una colección documental.

Exportar colección: permite exportar una colección documental que contenga documentos de archivos asociados, en un paquete compactado con extensión "zip" que contiene las descripciones de los documentos de archivos y sus representaciones digitales.

Compartir colección: permite seleccionar una colección documental que contenga documentos de archivos asociados y compartirla con cualquier otro usuario en el sistema.

Buscar colecciones compartidas: permite realizar una búsqueda de colecciones documentales que le han sido compartidas al usuario autenticado.

Listar documento de colección compartida: permite consultar los documentos que pertenecen a la colección compartida especificada.

Aceptar colección compartida: permite luego de visualizar los documentos de la colección compartida aceptarla, para que forme parte de la lista de colecciones pertenecientes al usuario autenticado.

Rechazar colección compartida: permite luego de visualizar los documentos de una colección compartida rechazarla, eliminándola de la lista de colecciones compartidas.

Mezclar documentos de colección compartida: permite luego de visualizar los documentos de la colección compartida mezclar dichos documentos con los documentos de la colección del usuario autenticado que tiene ese mismo nombre.

Reemplazar documentos de colección compartida: permite luego de visualizar los documentos de la colección compartida reemplazar los documentos que contiene la colección del usuario autenticado que tiene ese mismo nombre, por los documentos de la colección compartida.

Con la idea y con el fin de cumplir todos los requerimientos funcionales y no funcionales que requiere la Exportación de colecciones documentales, se ha decidido construir un módulo, con el uso del framework CodeIgniter y del servidor de base de datos MySQL.

Se identificó que en la actualidad no existe una norma o estándar que permita exportar colecciones

documentales que contengan las descripciones de los documentos de archivos asociados a la colección y sus representaciones digitales. Por ello se propone como solución en la presente investigación exportar las colecciones documentales del Sistema ArchiVenHIS en un fichero compactado con extensión “.zip” nombrado por el código y el nombre de la colección exportada, el cual contiene las descripciones de los documentos de archivos utilizando el estándar EAD, así como las representaciones digitales asociadas a dichas descripciones.

Se propone para el paquete exportado la siguiente estructura: un directorio nombrado “ead”, un directorio nombrado “representaciones_digitales” y un fichero nombrado “manifest.xml” que permite verificar la integridad de la colección, el cual contiene un listado con cada uno de los ficheros que agrupa el compactado, especificando el nombre, la extensión, el tamaño de cada fichero y un número resumen como resultado de una función HASH11 aplicada a cada uno de ellos.

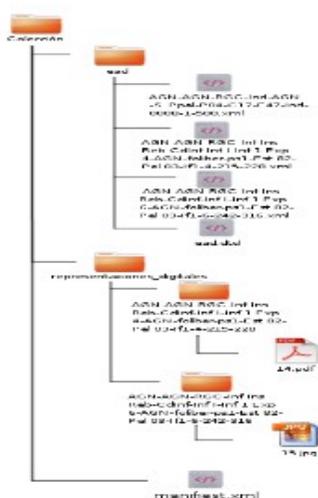


Figura 2. Salida del módulo exportación.

El directorio “ead” contiene ficheros con extensión “xml” de cada una de las descripciones de los documentos de archivos asociados a la colección, los cuales se identifican por el código de referencia de la descripción a la que corresponden; además, contendrá un fichero nombrado “ead.dtd”, el cual es la Definición de Tipo de Documento que permite validar las descripciones utilizando el estándar EAD.

La estructura básica de los instrumentos de descripción archivística codificada consiste en tres segmentos:

- **Cabecera EAD (EAD Header) <eadheader>**: contiene información similar a la de la portada de un texto impreso, pero relativa al documento electrónico, que identifica la versión electrónica del instrumento de descripción y puede también documentar las prácticas de descripción y codificación seguidas en su creación.
- **Material Inicial (Front Matter) <frontmatter>, opcional**: puede usarse para codificar estructuras como prefacios, dedicatorias u otro texto relativo a la creación, publicación o uso del instrumento de

descripción y para generar una portada adaptada a las necesidades o gustos locales.

- **Descripción Archivística (Archival Description) <archdesc>**: contiene la descripción archivística propiamente dicha y puede ser de un único nivel o contener una descripción multinivel con la descripción de componentes subordinados (las series de un fondo, las unidades documentales de éstas, en dos o más niveles).

El directorio “representaciones_digitales” contiene un directorio para cada una de las descripciones que tenga un archivo digital asociado, estos son identificados por el código de referencia de la descripción, de manera tal que exista una correspondencia entre una descripción y su archivo digital asociado, además dentro de estos directorios se encuentra el fichero digital correspondiente al documento original.

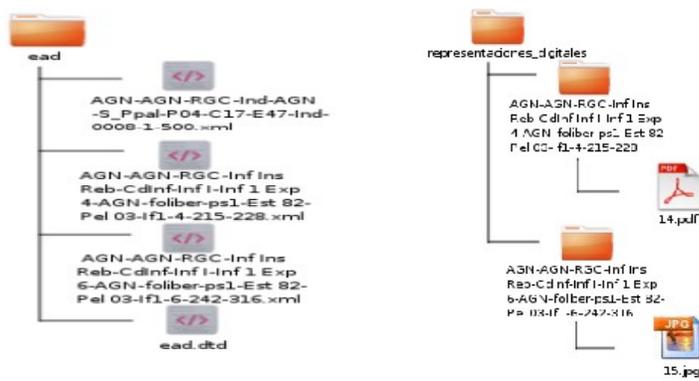


Figura 3. Salida del módulo exportación.

APORTE Y NOVEDAD

Se propone un formato para exportar colecciones documentales de un sistema de gestión de archivos, con sus representaciones digitales y las descripciones archivísticas de cada uno de los documentos asociados a la colección

CONCLUSIONES

Durante el desarrollo de la presente investigación se expuso la necesidad de implementar un módulo que permita exportar colecciones documentales del Sistema ArchiVenHIS, arribando a las siguientes conclusiones:

- Se identificó que internacionalmente para la descripción archivística se utiliza conjuntamente con normas particulares de cada país o institución la norma ISAD (G) y para el intercambio de información archivística entre sistemas informáticos se emplea el estándar EAD. Para el intercambio de colecciones documentales entre sistemas informáticos no se encontró ningún estándar que facilite esta actividad.

- Se identificó que existe un sistema nombrado Archon que permite la conformación de colecciones

documentales, sin embargo no se encontró ningún sistema que implemente su exportación; esto puede ser debido a la no existencia de un estándar para su empaquetado. Se realizó una propuesta para el empaquetado de colecciones de documentos junto con sus representaciones digitales.

➤ Con la implementación del módulo exportación de colecciones documentales para el sistema ArchiVenHIS se logra crear y conformar colecciones documentales, las cuales puede ser exportadas y consultadas fuera de AGN.

REFERENCIAS

1. **Consejo Internacional de Archivos.** Normal Internacional General de Descripción Archivística ISAD (G).
Madrid : s.n., 2000.
2. **Corral, Fernando Luis de Lanzas Sánchez del.** *Consejos Internacional de Archivos, ISAD (G) Norma Internacional General de Descripción Archivística.* Estocolmo, Suecia. : s.n., 19-22 de Septiembre de 1999, Vol. 122.
2. **Eito-Brun, Ricardo.** *E-LIS. E-prints in Library and Information Science. Lenguajes de marcado y archivos digitales.* [En línea] 2005. [Citado el: 07 de Octubre de 2011.] <http://eprints.rclis.org/handle/10760/13060>.
3. **Mena Mugica, Mayra.** *Gestión Documental y organización de archivos.* Ciudad de la Habana, Cuba : Félix Varela, 2005.