

Tipo de artículo: Artículo original

## Sistema informático para la digitalización del expediente académico del archivo histórico de la secretaría docente

### *Computer system for the digitization of the academic record of the historical archive of the teaching secretariat*

Haraid Martínez Cano<sup>1\*</sup> 

<sup>1</sup> Departamento Ciencias Básicas, Facultad 3, Universidad de las Ciencias Informáticas. [hmartinez@uci.cu](mailto:hmartinez@uci.cu)

\* Autor para correspondencia: [hmartinez@uci.cu](mailto:hmartinez@uci.cu)

#### Resumen

El archivo en una universidad es uno de los elementos que se obtiene como resultado del proceso de enseñanza, ya que respalda la evidencia de la trayectoria estudiantil de cualquier alumno que haya pasado por la institución. Hoy en día con los avances que muestran las Tecnologías de la Información y las Comunicaciones el proceso de conservación que requiere el archivo se logra mediante la digitalización de los documentos que se encuentran en él. Este proceso de digitalización es por lo general engorroso y propenso a errores por lo que es significativo el uso de herramientas informáticas que humanicen el trabajo. En el presente documento se propone la utilización de una solución informática que contribuya al proceso de digitalización y que de la posibilidad de organizar los documentos escaneados en la misma estructura en la que se encuentran en el archivo físico. Para el análisis de la documentación relacionada con el tema de investigación se utilizó el método analítico-sintético. Para construir la solución se utilizaron herramientas y tecnologías que utilizan licencia GNU<sup>1</sup> lo que garantiza soberanía tecnológica a la solución. Con el recurso puesto a disposición de la secretaría se contribuyó a la digitalización, conservación y preservación del expediente académico.

**Palabras clave:** archivo histórico; digitalización; expediente académico; secretaría docente; sistema informático.

#### Abstract

*The file in a university is one of the elements that is obtained as a result of the teaching process, since it supports the evidence of the student trajectory of any student who has passed through the institution. Nowadays, with the advances that Information and Communication Technologies show, the preservation process that the archive requires is achieved through the digitization of the documents found in it. This digitization process is generally cumbersome and prone to errors, which is why the use of computer tools that humanize the work is significant. This document proposes the use of a computer solution that contributes to the digitization process and that of the possibility of organizing the scanned documents in the same structure in which they are in the physical file. For the analysis of the documentation related to the research topic, the analytical-synthetic method was used. To build the solution, tools and technologies that use the GNU license were used, which guarantees technological sovereignty to the solution. The resource made available to the secretariat contributed to the digitization, conservation and preservation of the academic record*

**Keywords:** historical archive; digitization; academic record; teaching secretariat; computer system

<sup>1</sup> GNU es un sistema operativo de tipo Unix, lo cual significa que se trata de una colección de muchos programas: aplicaciones, bibliotecas, herramientas de desarrollo y hasta juegos. El nombre «GNU» es un acrónimo recursivo de «GNU No es Unix» GNU se usa generalmente con un núcleo llamado «Linux». (Free Software 2021)



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

**Recibido: 05/04/2021****Aceptado: 05/08/2021**

## Introducción

Del actuar de la Administración Pública deriva un patrimonio documental que constituye una parte esencial de la memoria histórica colectiva, al tiempo que informa la actuación pública de manera permanente, por lo que se debe prestar una especial atención al tratamiento, custodia y difusión de los documentos públicos. La Universidad de las Ciencias Informáticas (UCI) genera y recibe una considerable cantidad de documentación como consecuencia de las actividades necesarias para el cumplimiento de sus fines que son: el estudio, la docencia y la producción -investigación. Esta actividad administrativa se refleja en la documentación que genera que, puede tener otros valores como son: el fiscal, el legal y en su última fase de vida, el histórico, la memoria de esta institución.

Pero esta actividad no sólo tiene importancia en la esfera interna de la administración, también tiene una dimensión externa puesto que la administración universitaria no gestiona papel, gestiona información, y ésta tiene consecuencias en la órbita de derechos de las personas usuarias, derechos que hay que garantizar, también tiene consecuencias en la propia organización universitaria cuya actuación puede ser objeto de control y verificación.

Esta institución en la Secretaría General de la universidad lleva un expediente académico de cada estudiante por cada año académico. Este archivo tiene información valiosa que es preciso conservar y preservar durante ciclos de vida que puedan llegar a ser permanentes. En este contexto el archivo de la universidad se encuentra expuesto al deterioro que está condicionado por el tiempo de vida que tienen los documentos en formato duro. Además, algunos insectos como los ácaros atacan este tipo de material, y producen enfermedades en las personas que deben trabajar con estos documentos. Como alternativa y para lograr la conservación adecuada de estos documentos se utilizan herramientas informáticas que permiten la digitalización de este archivo.

Actualmente el proceso de digitalización se hace página a página del expediente de cada estudiante para que luego otro personal se encargue de organizarlo por el orden de cada página y por las carpetas de acuerdo a la estructura. Este proceso se hace sin tener en cuentas las normas definidas para la organización archivística y documental. Por otra parte, con frecuencia se incurren en errores como la duplicación de documentos, y que el documento escaneado si no se encuentra en buenas condiciones de conservación la digitalización obtenida no tiene la suficiente calidad para luego ser utilizado.



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**  
(CC BY 4.0)

Para el proceso de digitalización de archivos ya existen otras soluciones informáticas pero no cumplen en su totalidad con las necesidades de la secretaría docente de la Universidad, algunas de sus principales características se muestran más adelante en el documento y se argumenta porque no se pueden seleccionar como parte de la solución.

Ante estas deficiencias se evidencia la necesidad de desarrollar una solución informática que de la posibilidad de escanear el expediente de un estudiante y que la salida de este proceso de digitalización coincida con la estructura y directorio de organización que se utiliza en la secretaría. Que de manera automática corrija la imagen obtenida para lograr un resultado de mayor calidad visual y que contribuya a la búsqueda de cualquier expediente en tiempos de respuestas casi instantáneos. Por lo que el objetivo trazado es desarrollar una aplicación informática que contribuya al proceso de digitalización del expediente académico de los estudiantes del archivo histórico de la Secretaría General de la UCI.

## Materiales y métodos

Para guiar y facilitar el trabajo de los desarrolladores en la creación y documentación del software se utilizó la metodología ágil Programación Extrema (por sus siglas en inglés XP). Las metodologías ágiles combinan una filosofía y conjunto de directrices de desarrollo. La filosofía busca la satisfacción del cliente y la entrega temprana de software incremental, equipos de proyecto pequeños y con alta motivación, métodos informales, un mínimo de productos de trabajo de la ingeniería de software, y una simplicidad general de desarrollo. Las directrices de desarrollo resaltan la entrega sobre análisis y diseño (aunque estas actividades no se descartan) y la comunicación activa entre los desarrolladores y los clientes. XP usa un enfoque orientado a objetos como paradigma preferido de desarrollo, y engloba un conjunto de reglas y prácticas que ocurren en el contexto de cuatro actividades estructurales: planeación, diseño, codificación y pruebas. (Pressman 2010)

### Herramientas y tecnologías

Para el desarrollo de la aplicación  $\delta\alpha\rho\chi\epsilon\acute{\iota}\omega$ <sup>2</sup> se empleó el lenguaje Java en su versión 1.8 con la tecnología javaFX 2.2. Este es un lenguaje de programación y una plataforma informática comercializada por primera vez en 1995 por *Sun Microsystems*, es rápido, seguro y fiable.) (Oracle Corporation 2018b). JavaFX es una tecnología que amplía la potencia de Java permitiendo a los desarrolladores utilizar cualquier biblioteca de Java en aplicaciones JavaFX. De esta forma, los desarrolladores pueden ampliar sus capacidades en Java y utilizar la

---

<sup>2</sup> Su nombre son las siglas DARCHEIO (digitalización de archivos) escrito en griego haciendo honor a las primeras civilizaciones que crearon archivos documentales.



tecnología de presentación que JavaFX proporciona para crear atractivo visual. Como usuario, podrá ejecutar aplicaciones JavaFX en un explorador o arrastrarlos y soltarlos en el escritorio. (Oracle Corporation 2018a)

El entorno de desarrollo (por sus siglas en inglés IDE) utilizado fue NetBeans en su versión 8.2. Este software es un entorno de desarrollo escrito en JAVA pero puede servir para cualquier otro lenguaje de programación ya que posee un número importante de módulos que permite extender sus funcionalidades básicas, es un producto libre y gratuito sin restricciones de uso. (calendamaia 2014)

Para el tratado de imágenes se utilizó la herramienta Inkscape, la cual es un editor profesional de gráficos vectoriales para Windows, Mac OS X y Linux, libre y de código abierto (Freely 2021)

Como base de datos se utilizó SQLite en su versión 3.19.3. Este es un motor de base de datos de lenguaje de consulta estructurado (por sus siglas en inglés SQL). A diferencia de la mayoría de las otras bases de datos SQL, SQLite no tiene un proceso de servidor separado. SQLite lee y escribe directamente en archivos de disco duro. Contiene un único archivo de disco con una base de datos SQL completa con múltiples tablas, índices, disparadores y vistas. El formato de archivo de la base de datos es multiplataforma por lo que se puede copiar libremente una base de datos entre sistemas de 32 bits y 64 bits. (Oracle Corporation 2021)

## Resultados y discusión

En el mundo actualmente existen diferentes sistemas o herramientas, tanto de software libre<sup>3</sup> como propietarias<sup>4</sup>. Estas herramientas permiten dar solución a diversos problemas relacionados con la digitalización de documentos y gestión documental, reduciendo a su vez los asociados a la organización de la información dentro de la institución. A continuación, se expone algunos de estos sistemas:

- Internacionales
- Scan2PDF

La herramienta permite escanear documentos y guardarlos en formato PDF en el ordenador. También permite abrir imágenes que se tengan en el disco duro, y combinarlas con un documento recién escaneado, o reunir varios documentos o fotos en un solo PDF (Freeware OCR 2018). Es fácil de usar y la interfaz cuenta con un

---

<sup>3</sup> «Software libre» es el software que respeta la libertad de los usuarios y la comunidad. A grandes rasgos, significa que los usuarios tienen la libertad de ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software. Es decir, el «software libre» es una cuestión de libertad, no de precio. (Luis Miguel Arteaga 2018)

<sup>4</sup> Se le llama software propietario, no libre, privado o privativo al tipo de programas informáticas o aplicaciones en el que el usuario no puede acceder al código fuente o tiene un acceso restringido y, por tanto, se ve limitado en sus posibilidades de uso, modificación y redistribución. Este tipo de software se opone al más recientemente popularizado software libre, que permite que cualquiera lo modifique y lo redistribuya. (Bembibre 2009)



diseño sencillo. Presenta las siguientes características: Carga imágenes desde el disco duro. Escanea imágenes. Cambia el orden de las imágenes. Borra imágenes de la lista. Rota imágenes.

### **Nuxeo**

Nuxeo es un software que permite implementar con gran funcionalidad un repositorio documental corporativo. Su implementación es sencilla si lo que se quieren cubrir son necesidades no muy específicas («Productos de Nuxeo | Nuxeo» 2018). Esta solución te ofrece las siguientes características: Colaboración, flujos de trabajo, búsquedas eficientes. Flexibilidad, empleando una arquitectura basada en estándares. Robustez, utilizando Java para entornos Enterprise (J2EE)<sup>5</sup> entre otras tecnologías. Velocidad, gracias a la integración con el rápido motor de búsqueda Lucene. Capacidad de evolución en captura inteligente de documentos mediante Athento. Seguridad gracias a la implementación de estándares como SSL<sup>6</sup>, Single Sign On (SSO)<sup>7</sup> y facilitando el cumplimiento con normativas como la ISO 27.001 (Seguridad de la Información) o la Ley de Protección de Datos (LOPD).

### **Alfresco**

Es una solución versátil compatible con software tanto de la vertiente Microsoft, como de la rama Linux. Posibilita la creación y gestión de contenidos empresariales desde una gran cantidad de blogs y paquetes ofimáticos (Office y OpenOffice). Además, ofrece una gran variedad de herramientas colaborativas como calendarios individuales y de equipo, comentarios de actividades y tableros de discusión. Alfresco es para las empresas ante todo colaboración, pero también constituye un gestor documental. Su base de programación, junto a Nuxeo, es Java, lo que los convierte en soluciones multiplataforma adaptables a cualquier entorno («Alfresco» 2018). Entre sus principales características se encuentran: Facilidad de uso. Entorno web. Soporte para Gestión de Contenidos Empresariales (incluidas Gestión Documental y Gestión de Activos Digitales). Soporte para Gestión de Contenido Web (WCM). Soporte para Records Management. Soporte para Gestión de Contenido Social (Colaboración). Fácil despliegue y administración. Escalabilidad.

### **Athento – Software de Gestión Documental Inteligente**

---

<sup>5</sup> J2EE: Java 2 Enterprise Edition. Orientado a empresas y a la integración entre sistemas. Incluye soporte a tecnologías para internet. Su base es J2SE. (Rodríguez 2021)

<sup>6</sup> SSL es una tecnología estandarizada que permite cifrar el tráfico de datos entre un navegador web y un sitio web (o entre dos servidores web), protegiendo así la conexión. Esto impide que un hacker pueda ver o interceptar la información que se transmite de un punto a otro, y que puede incluir datos personales o financieros. (Alsina 2017)

<sup>7</sup> Single Sign On conocido también como SSO es una manera de acceder a múltiples aplicaciones interrelacionadas, aunque independientes, con las que el usuario solo tiene que iniciar sesión una única vez en lugar de introducir los datos de acceso individuales de cada software. («SSO» 2021)



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Es una aplicación informática que permite el manejo, gestión, conservación, publicación y trabajo sobre documentos electrónicos. Te permite ahorrar tiempo y esfuerzo en la gestión de documentos (Athento 2018). Athento es una suite compuesta por un conjunto de módulos que permiten cubrir el ciclo de vida de los documentos desde su captura hasta su archivo: Athento ECM: Gestor documental. Athento Smart Engine (SE): Captura de documentos. Athento Rhombus: Herramienta visual de parametrización del gestor documental. Todos los componentes de la suite de Athento están cobijados por la misma licencia de uso. Es decir, que el cliente contará con el uso de todos los módulos bajo una única suscripción.

Athento –*Smart Document Management*– permite a las empresas automatizar procesos relacionados con la captura, gestión, almacenamiento y distribución de documentos. A diferencia de otros sistemas, Athento permite a las compañías contar con un proceso de Enterprise Content Management integrado en sus diferentes fases (Athento 2018).

Athento ha sido desarrollado mediante módulos. Sus dos módulos fundamentales son el módulo de Captura y el módulo ECM. Estos dos módulos, integrados de forma nativa, pero independientes, cubren las fases de Captura, Almacenamiento, Gestión y parte de la Distribución (Athento 2018).

### **Kofax Express**

Facilita a todo tipo de usuarios la digitalización, organización y almacenamiento de documentos, a velocidades que aceleran drásticamente el procesamiento de lotes, tanto grandes como pequeños. Express es suficientemente fácil para los principiantes y suficientemente potente para los expertos (Kofax 2018)

### **SE Capture**

Es un software que simplifica el tratamiento de aspectos como la captura, procesamiento, validación y gestión de contenido. Además, notifica automáticamente a los equipos y departamentos sobre nuevos contenidos. A través de los recursos de procesamiento automático de imágenes, reconocimiento, clasificación, extracción de datos e indexación de documentos, las organizaciones pueden agregar valor y enriquecer el contenido, acelerando los procesos de captura, evitando errores y elevando los niveles de productividad (Capture 2021). Algunas de sus principales características son: Soporta escáner, cámaras digitales y otros dispositivos compatibles con el estándar TWAIN<sup>8</sup>. Clasifica documentos a través de categorías jerárquicas. Acelera los procesos de captura y reduce intervenciones manuales a través de la programación de los procesos. Mejora la productividad a través de la monitorización de carpetas o cuentas de e-mail, e importación automática del contenido. Permite establecer

---

<sup>8</sup> TWAIN (Technology Without An Interesting Name), Tecnología Sin Un Nombre Interesante. Este protocolo es el responsable de facilitar la comunicación entre el dispositivo de digitalización y el programa de retoque fotográfico. La interfaz desde la que ajustar todos los parámetros que intervienen en el proceso de digitalización de una imagen recibe el nombre de TWAIN. (García 2021)



procesos de captura con etapas como control de calidad, separación de documentos, reconocimiento de caracteres, verificación, indexación, entre otros, de acuerdo con el tipo de documento. Detecta y elimina automáticamente las páginas en blanco, para optimizar el espacio de almacenamiento. Permite priorizar actividades de captura, a través del acompañamiento de las actividades en día y en atraso. Captura un único tipo de documento, o un gran volumen de documentos diferentes en procesos complejos y multidepartamentales. Almacena archivos en banco de datos o en directorios controlados, aplicando mecanismos de compresión y protección por contraseña. Convierte las imágenes capturadas en imágenes (TIFF, JPEG, FIG) o PDF rastreable. Ofrece visualizador nativo embarcado en la solución para la visualización de documentos PDF, TIFF, JPEG, GIF, entre otros. Aplica controles de seguridad por usuario, equipo, categoría, documentos, entre otros. Localiza documentos a través de la aplicación de filtros por usuario, departamento, metadato, entre otros. Presenta informes de productividad para el acompañamiento de identificación de posibles cuellos de botella en los procesos de captura. Permite exportar informes para Excel para manipulación de las informaciones.

### **Dokmee**

Es una herramienta segura y fácil de usar. Sistema de gestión documental diseñado para una variedad de propósitos, incluyendo la captura de documentos y el almacenamiento, búsqueda y recuperación y compartición de archivos. Dokmee se adapta a cualquier modelo de negocio al maximizar la accesibilidad y funcionalidad en los repositorios de todos los tamaños, mientras aumenta la colaboración y la comunicación entre los usuarios. Con una interfaz fácil de usar disponible en 19 idiomas.(dokmee 2018)

Las integraciones entre Dokmee y otros programas, ambos basados en Windows y basada en la Web, pueden utilizar la búsqueda y recuperación, indexación, la importación y la visualización de los archivos. La barra de herramientas de fácil acceso de Microsoft Office le permite enviar archivos directamente desde Word, Excel y Outlook directamente en Dokmee. Con la Dokmee Virtual Printer, envía archivos PDF en Dokmee desde cualquier aplicación utilizando el botón de impresión. Utiliza una estructura de carpetas como Windows, de forma manual y automáticamente permite ordenar los archivos en un número ilimitado de carpetas y niveles. (dokmee 2018)

Dokmee Desktop es una solución de gestión documental independiente diseñada para la pequeña oficina o el usuario de oficina en casa. Este sistema completamente funcional permite el almacenamiento y recuperación de archivos electrónicos y mucho más con un fuerte Microsoft SQL back-end incluido gratis. Tasado agresivamente para el usuario individual, Dokmee Desktop es una excelente alternativa para el almacenamiento de su información en carpetas en su equipo(dokmee 2018).

Según (dokmee 2018) Algunas características que tiene esta herramienta son: Comparte archivos en su organización. Almacenar y gestionar cualquier tipo de archivo en Dokmee, con construido en capacidades de



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

visualización de PDF, TIFF, JPG, PNG, BMP, GIF, DWG, DXF, MSG, EML, y mucho más. Añadir anotaciones, márgenes y notas a los archivos. Con un visor de Microsoft Office colaborativo, abrir y editar archivos de Word, Excel, PowerPoint y Visio. Varios usuarios pueden ver los archivos al mismo tiempo, junto con el control de versiones para ediciones y revisiones.

## Nacionales

### DigiPyrus

El sistema de Digitalización DigiPyrus, desarrollado por el Centro de Gobierno Electrónico (CEGEL) de la Universidad de las Ciencias Informáticas (UCI), permite digitalizar los Tomos de los Registros y Notarías de la República Bolivariana de Venezuela, contribuyendo a agilizar el acceso a la información que a diario se tramita en cada una de sus oficinas. Este sistema surge debido a la demanda de una gestión más eficiente y segura de los trámites que se realizan con los documentos archivados en las Oficinas de los Registros y Notarías Públicas. De ahí, que se haya propuesto presentar una solución que cumpla con el objetivo de dotar a esta entidad de un sistema que contenga las funciones, para realizar el proceso de digitalización de todos los documentos archivados y de esta forma obtener el fondo digital de cada una de estas (MONTES 2007).

Los procesos que se informatizan (MONTES 2007) mediante este sistema son: Recepción de Documentos: Verificación y registro de los tomos provenientes del Almacén Temporal. Preparación de documentos: Ejecución de las operaciones establecidas para garantizar que los tomos cumplan con los parámetros necesarios para el proceso de digitalización. Escaneo de Documentos: Conversión de los tomos físicos en objetos digitales multipáginas. Control de la Calidad: Inspeccionar el 100% del objeto digital multipáginas proveniente del área de Digitalización de Documentos. Encuadernación de Documentos: Recuperación y preservación de las condiciones iniciales de los tomos digitalizados. Devolución de Documentos: Registro de salida de los tomos que han sido digitalizados en el proceso.

### DigiDaP

DigiDAP forma parte de la Solución Tecnológica Integral para la automatización y modernización de la División de Antecedentes Penales de la República Bolivariana de Venezuela, como el subsistema que informatiza los procesos del Centro de Digitalización para el fondo documental de dicha institución. Su objetivo fundamental es garantizar la obtención de objetos digitales con valor legal a partir de la digitalización del fondo para mejorar los servicios de inscripción y certificación de antecedentes penales a los ciudadanos de la nación venezolana (González y Suárez 2011)

En el modelo de referencia para la digitalización de fondos documentales de antecedentes penales (González y Suárez 2011) se tiene: Recepción y Devolución de Documentos: Recepcionar expedientes al centro de digitalización, devolver desde este al archivo y eliminar expedientes que por errores se decidan quitar del



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)



proceso. Preparación de Documentos: Preparar los expedientes y garantizar que tengan las condiciones necesarias para pasar por el escáner en la próxima área, evaluar la calidad de las unidades documentales para determinar si proceden o no en el proceso y registrar los folios a digitalizar de cada una y emitir notas de preparación en caso de que sea necesario para esclarecer la omisión de folios a digitalizar. Digitalización de Documentos: Digitalizar las unidades documentales contenidas en los expedientes, rectificar errores en los trámites que hayan sido regresados en el proceso y mejorar la calidad visual de los documentos digitales. Control de Calidad: Revisar legalmente los trámites realizados en el centro de Digitalización, emitir la respuesta correspondiente del trámite e imprimir salidas del sistema. Otorgamiento: Otorgar trámites revisados legalmente, certificando la calidad del proceso realizado, así como otorgar firma manuscrita a las salidas impresas del sistema y la firma digital a los documentos digitales, obtener reportes estadísticos sobre el proceso y localizar trámites dentro del proceso y consultar el estado de los mismos. Encuadernación de Documentos: Encuadernar y foliar las unidades documentales dentro de los expedientes para su asentamiento en archivo.

### **CDA**

CDA es la solución tecnológica desarrollada para el Centro de Digitalización de Alfabéticas para el Servicio Administrativo de Identificación, Migración y Extranjería. Para su desarrollo se utilizó como marco de trabajo el utilizado por la solución DigiPyrus por lo que las características arquitectónicas y tecnologías de desarrollo utilizadas son similares. Su objetivo fundamental es extraer tarjetas alfabéticas para convertirlas en metadatos. Esta solución está compuesta por varios módulos, algunos de estos son (MONTES 2007): Almacén Temporal: permite la gestión de inventario de los recursos que entran y salen del Centro de Digitalización, así como del control de recibo y despacho de las Tarjetas de Alfabética desde las Oficinas de Identificación hacia cada Centro de Digitalización. Preparación: permite la creación de los lotes de documentos en el sistema, se registran los documentos rechazados por malas condiciones para entrar en el proceso de digitalización y se realiza la corrección de los documentos rechazados por mala preparación o códigos de barra ilegible. Digitalización: permite la digitalización de los lotes de documentos o de documentos rechazados. Firma digital: permite la emisión de la firma digital de los documentos, certificando que la información obtenida es verídica.

### **Conclusiones del análisis de los sistemas existentes**

En la **Tabla 1** de resumen se muestra una comparación entre los sistemas anteriormente mencionados, reflejando sus ventajas y deficiencias dando paso a la argumentación de la idea del desarrollo de un nuevo sistema. Para la realización de esta tabla se tuvo en cuenta la licencia de pago, si organizaba la información, si es un sistema multiplataforma y si digitaliza o no.



**Tabla 1:** Tabla comparativa de sistemas homólogos

Software	Indicadores			
	Organización de la Información	Digitalización	Multiplataforma	Licencia de pago
<b>Scan2PDF</b>	No	Sí	No	No
<b>Nuxeo</b>	Sí	No	Sí	Sí
<b>Alfresco</b>	Sí	Sí	Sí	Sí
<b>Athento</b>	Sí	No	Sí	Sí
<b>Kofax Express</b>	Sí	Sí	No	Sí
<b>SE Capture</b>	Sí	Sí	No	Sí
<b>Dokmee</b>	Sí	No	Sí	Sí
<b>DigiPyrus</b>	Sí	Sí	Sí	Sí
<b>DigiDAP</b>	Sí	Sí	Sí	Sí
<b>CDA</b>	Sí	Sí	Sí	Sí

Fuente: Creación propia

El análisis guiado por la tabla anterior permitió conocer la existencia de sistemas de digitalización en Cuba y en el mundo, desarrollados en distintos lenguajes y con el mismo propósito. Sobre el estudio de las aplicaciones existentes a nivel internacional se tiene que no resultan soluciones factibles a tener en cuenta, pues fueron desarrolladas con software propietario lo que implica gastos muy elevados en licencia y mantenimiento. Aunque Scan2PDF no es privativo tampoco se puede considerar para la solución del problema, puesto que no es un sistema multiplataforma y no organiza la información que es una de las principales funcionalidades a tener en cuenta.

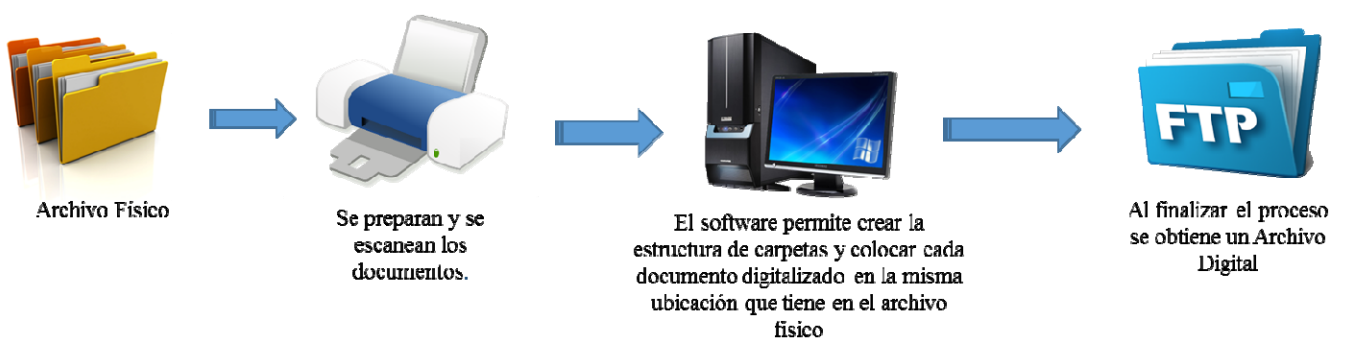
Sobre el estudio de las herramientas a nivel nacional, estas fueron desarrolladas con software propietario, por lo que no pueden ser usadas para la solución del problema expuesto en la investigación. Sin embargo, en un análisis a profundidad fueron identificadas un conjunto de características a tener en cuenta en la solución a construir como son: la adaptabilidad a cualquier entorno, la edición y clasificación de documentos digitales a través de categorías jerárquicas, la conversión de tomos físicos en objetos digitales multipáginas y la posibilidad de guardar archivos en formato PDF en directorios controlados.

### Solución informática



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

El software  $\delta\alpha\rho\chi\epsilon\iota$  tiene como meta llevar un archivo físico a un archivo digital, proporcionar mayor organización y control de la información digitalizada mediante normas y estándares para la gestión documental. Un archivo físico consta de un conjunto de documentos de diferentes tipos que se encuentra almacenados en una estructura de carpetas que tienen un orden lógico, por lo tanto, el archivo digital debe mantener dichas características. A continuación, en la figura 1 se ejemplifica el proceso de digitalización de un archivo con el software.



**Figure 1:** Diseño de la solución

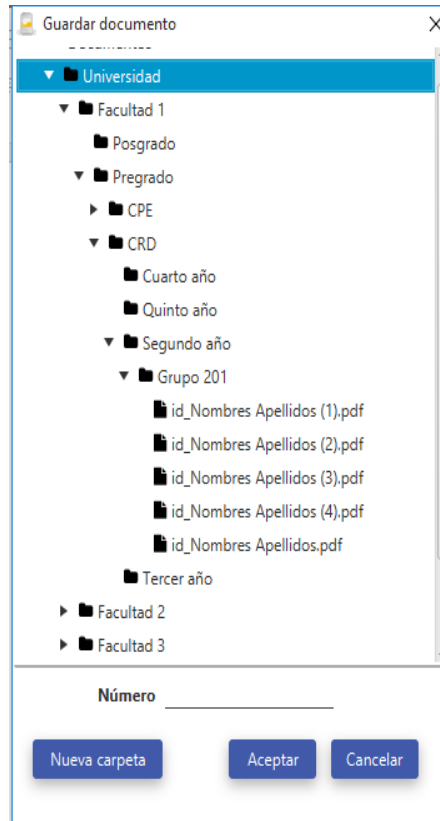
Teniendo en cuenta el proceso anterior se pueden definir las siguientes etapas:

**Preparación de los documentos a digitalizar:** es necesario preparar los documentos a digitalizar ya que no deben perder en ningún momento de la manipulación la estructura y el orden que tienen en el archivo físico. Identificar los que se encuentran en mejores condiciones de conservación para que sean los primeros en ser digitalizados y de esta manera ganar en rapidez. Es preciso tener en cuenta el grado de deterioro que tengan los documentos y hay que definir si se escanean en un escáner plano o en otro tipo de escáner. El uso del tracto masivo con un documento en mal estado puede llevar a su pérdida. Los documentos pueden estar presillados, engargolados o encuadernados, por tanto, hay que separarlos por páginas para poder digitalizarlos.

**Crear estructura de carpetas:** a través del software realizado se crean las estructuras de carpetas en un FTP<sup>9</sup> previamente creado. (ver **Figure 2**)

<sup>9</sup> Protocolos de la red Internet, concretamente significa File Transfer Protocol (Protocolo de Transferencia de Ficheros) y es el ideal para transferir grandes bloques de datos por la red. Su comportamiento está definido por la recomendación RFC 959. (Alsina 2016)





**Figure 2:** Estructura de carpetas en la que se almacena los documentos escaneados

**Digitalizar documentos:** se escanean los documentos a través del uso de un escáner y mediante el software se revisa cada imagen del documento para garantizar la fidelidad al documento original, se convierte a PDF<sup>10</sup> y se coloca en la estructura de carpeta previamente creada. (Ver Figure 3)

<sup>10</sup> PDF (sigla del inglés Portable Document Format, «formato de documento portátil») es un formato de almacenamiento para documentos digitales independiente de plataformas de software o hardware. Este formato es de tipo compuesto (imagen vectorial, mapa de bits y texto). (Cano 2014)



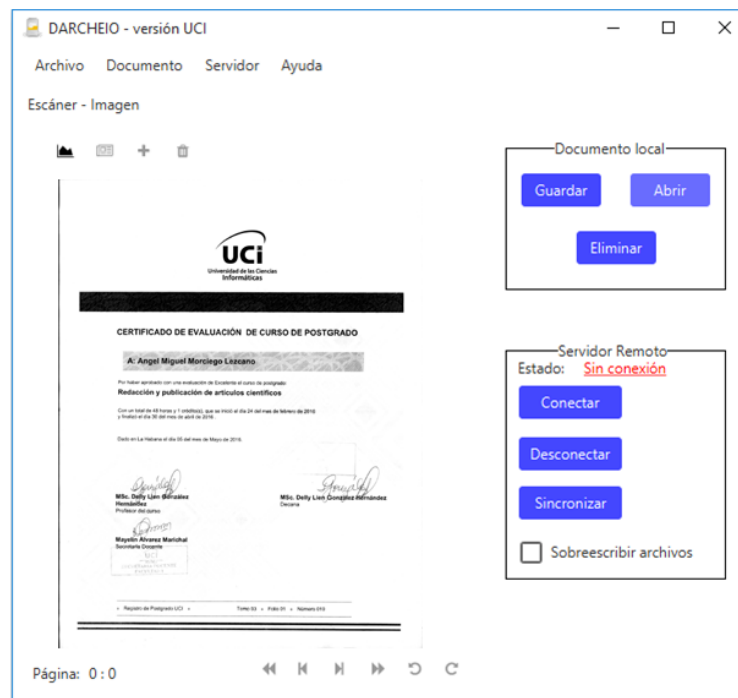


Figure 3: Interfaz principal de la aplicación, en la que se muestra un documento escaneado.

**Gestión documental:** los documentos digitalizados se archivan en el directorio local, de acuerdo con la clasificación definida en la norma Catalogación legible por máquina 21 (por sus siglas en inglés MARC) (SPEDALIERI 2001)

**Subir documentos al servidor remoto:** se sube el archivo digital creado en el ordenador local hacia el servidor remoto FTP para ser utilizado posteriormente por otros sistemas.

**Aspecto de seguridad que brinda el software:**

- Es importante tener en cuenta aspectos de seguridad en el software para que la información no se vea comprometida.
- El software es instalado en una PC a la cual solo puede acceder un único usuario, y el sistema requiere de autenticación de usuario para trabajar en él.
- El acceso al servidor se realiza utilizando un FTP seguro (sftp), basado en el protocolo SSH<sup>11</sup>, los datos viajan de manera segura desde la estación de trabajo donde se encuentra la aplicación hacia el servidor.
- A medida que los documentos digitales van siendo transferidos al FTP son borrados automáticamente de los temporales en la PC evitando que después se quede almacenada la información y sea propagada.

<sup>11</sup> SSH (Secure SHell, en español: intérprete de órdenes seguro) es el nombre de un protocolo y del programa que lo implementa, y sirve para acceder a máquinas remotas a través de una red (Alsina 2017)



Esta obra está bajo una licencia Creative Commons de tipo Atribución 4.0 Internacional (CC BY 4.0)

### **Aporte económico y social**

El sistema para la digitalización del archivo de la secretaría permite a la UCI preservar los documentos de una manera más eficiente además de que da la posibilidad de una mayor seguridad y una longevidad casi infinita. Posibilita gestionar los documentos según las normas y estándares de la gestión documental para poseer una mejor integración con otros sistemas de búsqueda de información.

δαρχείο ha permitido los siguientes beneficios:

✓ **Beneficios tangibles:**

- Disminución en el uso de papel e impresiones.
- Disminución en el costo de almacenamiento.
- Reducción de copias y copadoras (mantenimiento, consumibles).
- Disminución del personal dedicado a archivar y buscar información.
- Disminución en mantenimiento de archivos físicos.

✓ **Beneficios intangibles:**

- Más tiempo disponible para trabajar y menos tiempo perdido en archivar.
- Descentralización de la información.
- Acceso total y rápido a la información.
- Seguridad en los archivos.
- Mejor uso de la información para la toma de decisiones.
- Aprovechamiento de espacios usados en archivos (costo de oportunidad).
- Mayor productividad del personal.
- No más pérdida o traspapelado de documentos.
- Seguridad en encontrar toda la información archivada.

Actualmente el software se encuentra en uso en la secretaría docente general de la UCI donde se están digitalizando los expedientes de los estudiantes de todos los estudiantes. Es un proceso que tomará algún tiempo porque solamente hay dos personas en la secretaría y se deben procesar los expedientes archivados y los nuevos que se van recibiendo.



## Conclusiones

El presente artículo concluye que el software desarrollado contribuye a la digitalización y conservación del expediente académico en el archive histórico de la secretaría docente de los estudiantes de la Universidad de las Ciencias Informáticas. También se pudo demostrar que al digitalizar un documento este puede ser utilizado por varias personas a la vez lo que hace que la información se encuentre más accesible. Utilizando las bondades que posibilita Internet es posible poner a disposición desde la red de la institución y a las personas autorizadas los archivos durante las 24 horas del día los 7 días de la semana. Por otro lado, y como parte de la acreditación necesaria para cada institución, la herramienta construida favorece significativamente el proceso de digitalización de evidencias, elemento crucial para la revisión por parte de la Junta de Acreditación Nacional (JAN).

## Conflictos de intereses

El autor de la presente investigación declara que no poseen conflictos de intereses.

## Contribución de los autores

1. Conceptualización: Haraid Martínez Cano
2. Curación de datos: Haraid Martínez Cano
3. Análisis formal: Haraid Martínez Cano
4. Adquisición de fondos: Haraid Martínez Cano
5. Investigación: Haraid Martínez Cano
6. Metodología: Haraid Martínez Cano
7. Administración del proyecto: Haraid Martínez Cano
8. Recursos: Haraid Martínez Cano
9. Software: Haraid Martínez Cano
10. Supervisión: Haraid Martínez Cano
11. Validación: Haraid Martínez Cano
12. Visualización: Haraid Martínez Cano
13. Redacción – borrador original: Haraid Martínez Cano
14. Redacción – revisión y edición: Haraid Martínez Cano



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**  
(CC BY 4.0)

## Financiamiento

La presente investigación ha sido financiada con recursos propios de los autores.

## Referencias

- Alfresco. *Alfresco* [en línea], 2018. [Consulta: 1 noviembre 2018]. Disponible en: <https://www.itop.es/alfresco.html>.
- ALSINA, G., 2016. Definición de FTP. *Definición ABC* [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.definicionabc.com/tecnologia/ftp.php>.
- ALSINA, G., 2017. Definición de SSL, TLS y SSH. *Definición ABC* [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.definicionabc.com/tecnologia/ssl-tls-ssh.php>.
- ATHENTO, 2018. Software de Gestión Documental - Athento. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <http://www.athento.com/es/gestion-documental-inteligente/>.
- BEMBIBRE, V., 2009. Definición de Software propietario. *Definición ABC* [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.definicionabc.com/tecnologia/software-propietario.php>.
- CALENDAMAIA, 2014. NetBeans. *Genbeta* [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.genbeta.com/desarrollo/netbeans-1>.
- CANO, F., 2014. Definición de PDF. *Definición ABC* [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.definicionabc.com/tecnologia/pdf.php>.
- DOKMEE, 2018. Software de Gestión Documental | Gestión de Contenido Empresarial. [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://dokmee.com/es/>.
- FIDEL ALEJANDRO CEBRIÁN LA ROSA, J.S.M., 2014. *Plataforma para la construcción de sistemas de digitalización de documentos (DigiPRO)*. Tesis de grado. S.l.: UCI. TD\_07335\_14.pdf
- FREE SOFTWARE, 2021. El sistema operativo GNU y el movimiento del software libre. *El sistema operativo GNU* [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.gnu.org/home.es.html>.
- FREELY, D., 2021. Draw Freely | Inkscape : Inkscape es un editor profesional de gráficos vectoriales para Linux, Windows y macOS. Es libre y de código abierto. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://inkscape.org/es/>.
- FREWARE OCR, 2018. Freeware OCR Scanner Software Scan2PDF (Scan To PDF). *Freeware OCR Scanner Software Scan2PDF (Scan To PDF)* [en línea]. [Consulta: 27 julio 2021]. Disponible en: <http://www.scantopdf.org/>.



Esta obra está bajo una licencia **Creative Commons** de tipo **Atribución 4.0 Internacional** (CC BY 4.0)



- GARCIA, D.E., 2021. TWAIN - GlosarioIT: Glosario Informático. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.glosarioit.com/TWAIN>.
- GONZÁLEZ, S. y SUÁREZ, R., 2011. *Diseño e Implementación de los módulos de Preparación de Documentos, Digitalización de Documentos y Asociación de Metadatos del Centro de Digitalización para la División de Antecedentes Penales*. Tesis de grado. S.l.: Universidad de las Ciencias Informáticas. [http://repositorio.uci.cu/jspui/handle/ident/TD\\_04302\\_11](http://repositorio.uci.cu/jspui/handle/ident/TD_04302_11).
- KOFAX, 2018. Kofax Express | Kofax. *Kofax Express | Kofax*. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.kofax.com/Products/InformationCapture/Express/Features>.
- MARTÍNEZ, A.G., 2013. *Digitalización de Documentos*. Tesis de grado. S.l.: s.n.
- MONTES, L.R., 2007. *Análisis y Modelado del sistema para el proceso de digitalización de documentos de Registros y Notarias del MPPRIJ de la República Bolivariana de Venezuela*. 2007. S.l.: s.n.
- NORMAS MARC: Oficina de Desarrollo de Redes y Normas MARC. [en línea], 2021. Biblioteca del Congreso.: [Consulta: 27 julio 2021]. Disponible en: <https://www.loc.gov/marc/marcspa.html>.
- NUXEO, 2018. Productos de Nuxeo | Nuxeo. *Productos de Nuxeo | Nuxeo*. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.nuxeo.com/es/products/>.
- ORACLE CORPORATION, 2018a. Preguntas frecuentes de JavaFX. [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.java.com/es/download/faq/javafx.xml>.
- ORACLE CORPORATION, 2018b. ¿Qué es Java y para qué es necesario? [en línea]. [Consulta: 27 julio 2021]. Disponible en: <https://www.java.com/es/download/faq/javafx.xml>.
- ORACLE CORPORATION, 2021. SQLite Release 3.19.3 On 2017-06-08. [en línea]. [Consulta: 27 julio 2021]. Disponible en: [https://www.sqlite.org/releaselog/3\\_19\\_3.html](https://www.sqlite.org/releaselog/3_19_3.html).
- PRESSMAN, R.S., 2002. *Ingeniería del software. Un enfoque práctico*. 6ta Edición. S.l.: McGraw-Hill. ISBN 978-970-10-5473-4.
- PRESSMAN, R.S., 2010. *Software Engineering. A Practitioner's Approach*. 7ma Edición. New York: McGraw-Hill. ISBN 978-0-07-337597-7.
- ¿QUÉ SON SSL, TLS Y HTTPS...? *DigiCert* [en línea], 2021. [Consulta: 27 julio 2021]. Disponible en: </es/what-is-ssl-tls-https/>.
- RODRÍGUEZ, A., 2021. Versiones y distribuciones Java: J2SE, J2EE, J2ME. ¿Java 7, Java 8, Java9..., cuál es mejor? (CU00606B). [en línea]. [Consulta: 27 julio 2021]. Disponible en: [https://www.aprenderaprogramar.com/index.php?option=com\\_content&view=article&id=377:versiones-y-distribuciones-java-j2se-j2ee-j2me-ijava-7-java-8-java9-cual-es-mejor-cu00606b&catid=68&Itemid=188](https://www.aprenderaprogramar.com/index.php?option=com_content&view=article&id=377:versiones-y-distribuciones-java-j2se-j2ee-j2me-ijava-7-java-8-java9-cual-es-mejor-cu00606b&catid=68&Itemid=188).



- CAPTURE, 2021. Software para Captura e Escaneo de Documentos | SoftExpert Captura. [en línea]. [Consulta: 28 julio 2021]. Disponible en: <https://www.softexpert.com/es/produto/captura-documentos/>.
- SPEDALIERI, G., 2001. *Manual de MARC : cómo interpretar y usar registros MARC*. 2001. S.l.: s.n. 109-113
- SSO: un solo inicio de sesión para todo. *IONOS Digitalguide* [en línea], 2021. [Consulta: 27 julio 2021]. Disponible en: <https://www.ionos.es/digitalguide/servidores/herramientas/single-sign-on/>.



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)