

Digitalización por Lotes y su relación con la Captura de Documentos

Introducción

La necesidad de organizar, indexar y archivar en un sistema de Gestión Electrónica Documental los documentos que se encuentran en formato papel, y hacerlo de un modo rápido y eficiente, se convierte a menudo en un extraordinario reto. La implementación de software adecuado a esta tarea hará que podamos abordarlo de forma efectiva.

La digitalización por lotes es el proceso que se usa habitualmente para convertir el papel en documentos digitales. Hacer un buen uso de este concepto puede incrementar enormemente la productividad y asegurar el éxito en la puesta en marcha del sistema documental. Este estudio perfilará la metodología de la digitalización por lotes para ayudar a entender el proceso de conversión digital.

Esquema de la captura de documentos

Tanto si le prestamos atención como si no, cuesta mucho dinero gestionar los documentos con los que trabajamos a diario. El costo de cada documento podría dividirse en estas cinco funciones primarias: creación, captura, circulación o flujo, almacenamiento y recuperación. El tratamiento efectivo de estas funciones puede ahorrar a su negocio miles de euros en pérdida de productividad, de espacio de almacenamiento y de eficiencia profesional.

La captura de la documentación es una parte integral del problema del papel al que debemos enfrentarnos. ¿En qué consiste pues la captura de la documentación? Dicho simplemente, se trata de convertir un formato "papel" a un formato "electrónico". Entonces, ¿cómo capturo mis documentos en papel? Lo primero que nos viene a la mente es digitalizar o escanear. Para muchos de nosotros, este sería el primer paso lógico porque es lo que vemos a diario: montones de papel cubriendo cada espacio libre que tenemos disponible en nuestras oficinas.

La captura de documentos se realiza generalmente mediante un dispositivo de entrada llamado escáner, una aplicación de software que ayuda a clasificar e indexar el papel y una cantidad de espacio en disco para almacenar los nuevos documentos electrónicos.

Superada la simple captura de documentos, se pueden hacer muchas cosas para automatizar estos procesos. Reconocimiento de códigos de barras, búsquedas en bases de datos externas, proceso de formularios y OCR son métodos avanzados de captura que pueden aumentar en gran medida la productividad. Encontrar un proveedor experto que ayude a analizar el proceso de captura es crítico para poder implementar una solución con éxito.

¿Cómo empezar entonces? Lo primero que debemos hacer es buscar los productos que nos ayuden a resolver el problema. Después, asegurarse de que son fáciles de usar y de mantener, y que proporcionan las herramientas necesarias para cumplir con los objetivos marcados. Es mucho más importante encontrar un buen colaborador que entienda nuestro problema y que pueda aplicar efectivamente la tecnología disponible, que comprar simplemente hardware y software y esperar a que funcione.

La definición del proceso

La digitalización por lotes se divide principalmente en dos categorías: procesos por lotes automáticos y procesos por lotes manuales. Las dos categorías tienen diferentes criterios y definiciones. Cada una de ellas tiene su espacio en el mundo de la digitalización, y puede complementar a la otra. De hecho, a veces es conveniente usar una combinación de ambas para ofrecer la mejor solución de digitalización.

En un documento aparte http://www.sdm.es/ficheros/estudio_reto_gestion_documental.pdf definíamos las cinco funciones primarias de un proceso de captura:

1. Preparación de la documentación
2. Digitalización
3. Control de calidad
4. Indexación y clasificación
5. Migración de los documentos a una solución documental

La preparación y la indexación fueron tratadas en detalle en dicho documento. El presente documento abundará más en la digitalización, el control de calidad y la migración al gestor documental.

Digitalización

El proceso de digitalización consiste en la captura de los documentos que se encuentran en formato papel. Al definir este proceso deben tenerse en cuenta los siguientes aspectos:

1. Conseguir el escáner adecuado para el trabajo, basándose en criterios como tipos de documento y requerimientos de proceso. Sin el hardware adecuado, el proceso podría llegar a ser frustrante y hacerle perder su valioso tiempo.
2. Seleccionar el software de captura apropiado, tras revisar los tipos de documento a digitalizar, los índices a capturar y las necesidades de obtención de los ficheros de salida. Asegúrese de cumplir los requisitos marcados a priori.
3. Decida si alguno de los índices puede automatizarse con OCR zonal, reconocimiento de código de barras, consultas de bases de datos externas, etc.
4. Fiabilidad, disponibilidad de piezas y soporte, calidad de imagen y capacidad de alimentación son factores a tener en cuenta para elegir un buen escáner.

Control de calidad

Una etapa clave del proceso por lotes es su capacidad de proporcionar un control de calidad de los documentos y de los metadatos capturados. Tenga en mente que, una vez digitalizado el documento, el original ya no va a estar disponible.

La imagen electrónica debe por consiguiente ser considerada como algo vital, ya que se convertirá en el documento "original". El control de calidad se necesita para revisar los documentos digitalizados y así asegurarse de que cumplen el estándar del documento original. Una puntualización interesante a hacer aquí es que, en la mayor parte de los casos, la calidad de la imagen digitalizada supera al original, debido a las posibilidades de proceso para mejorar la imagen que incorpora el software de captura.

Las características a tener en cuenta en un buen modulo de control de calidad son:

1. Que disponga de opciones como añadir, insertar, reemplazar, girar, voltear y borrar páginas de un documento.
2. Que permita la edición de las imágenes, con opciones como enderezar, borrar bordes negros, cortar, recortar, cambiar la orientación de la página, limpiar puntos negros, etc.
3. Que admita la revisión de imágenes automática, vistas en miniatura, reordenación de páginas y zoom de las imágenes.
4. Que simplifique la corrección y la validación de los datos indexados.

Este conjunto de características proporcionarán al personal de control de calidad la posibilidad de asegurar la fiabilidad de los datos y una buena calidad de las imágenes, asegurando de ese modo un buen documento electrónico, incluso mejor que el "original".

Exportación de las imágenes

Una vez digitalizadas las imágenes, hecho el control de calidad y finalizado el proceso de indexación, las imágenes están preparadas para su almacenamiento en el sistema de Gestión Electrónica Documental.

El proceso de captura debería incluir un mecanismo para enviar las imágenes al sistema con la categorización adecuada, tal como se completó en el proceso de indexación. El proceso por lotes definirá el formato de salida del documento y la descripción de su título, facilitando su migración al correspondiente repositorio de datos del sistema documental. Sin una salida apropiada del software, los documentos pueden ser mal archivados, provocando que su recuperación se convierta en un asunto complicado, lo que negaría el principio que justifica un sistema de Gestión Electrónica Documental, que es facilitar la consulta y la recuperación de los documentos.

Indexación manual vs proceso automático

Son los dos modos principales de captura de los datos en el proceso de documentos por lotes. Como ya se ha mencionado, en muchas ocasiones se emplea una combinación de ambas técnicas de indexación, manuales y automáticas.

Definimos normalmente el proceso de indexación manual como aquel que implica el teclado de los índices por un operador que visualiza una imagen en la pantalla. Este es el método más lento de un proceso por lotes, aunque en algunos casos, este es el único posible. Una consideración a tener en cuenta: a veces es más rápido usar este método manual que gastar un tiempo adicional en la etapa de preparar los documentos colocando códigos de barras y configurando consultas a bases de datos externas que automatice el proceso de indexación.

Al considerar el proceso manual, es importante chequear la cantidad de datos requeridos en la indexación y el número de índices que se necesitan. Una buena práctica consiste en mantener esta cantidad en el mínimo, de modo que se pueda realizar tan rápido como sea posible. Limitar el número de campos a capturar es crítico en el proceso de indexación.

Existen herramientas en la mayoría de los productos para agilizar los procesos de indexación manual. Algunas de esas funciones son: indexación en base a una plantilla, uso de un desplegable con los valores posibles de un campo para que el usuario lo seleccione, campos "preformateados" con parte de los datos ya introducidos, campos con máscaras del tipo MM/DD/AAAA e indexación asistida por OCR.

El proceso automático puede incrementar en gran medida el rendimiento de documentos en el proceso de las imágenes. Cuanto más se pueda automatizar, mejor rendimiento obtendremos. Las características del software que aceleran el proceso automático son el OCR zonal, el reconocimiento de códigos de barras y las consultas de bases de datos externas. Al utilizar alguna de ellas, o todas estas características juntas, se presentan situaciones en las que el proceso de indexación se puede automatizar por completo.

Conclusión

Cuando se trata de procesar una gran cantidad de documentos en un corto espacio de tiempo, la digitalización por lotes en el proceso de captura de documentos es verdaderamente útil. También es muy importante tener claros los tipos de documento que nos vamos a encontrar, para poder decidirse por el hardware y software de digitalización más adecuados, y decidirse por los protocolos de entrada de datos manuales y automáticos más eficaces.

Mediante la implementación de estas técnicas se pueden digitalizar a diario desde varios cientos de páginas a varios miles en una única estación de captura. Es absolutamente crítico contar con el apoyo de un experto para que le guíe en la definición del proceso de digitalización. El simple hecho de seguir estas reglas le asegurará una solución satisfactoria.