

OFIMATICA Y DOCUMENTACION: LOS SISTEMAS DE GESTION INTEGRADA DE LA DOCUMENTACION

Carlota Bustelo Ruesta
Socia-Directora
Gabinete de Asesores Documentalistas S.A.

1. INTRODUCCION

En toda oficina, organización o empresa se producen y reciben muchos documentos diariamente. En ellos se contiene toda la información necesaria para el desarrollo de la propia actividad y por lo tanto toda organización de ellas genera un sistema, aunque sea inconsciente, para la gestión de los documentos. Si entendemos por "Gestión de documentos" todo el sistema administrativo por el que una organización controla la creación, distribución, archivado, recuperación, almacenamiento y expurgo de los documentos que son producidos o recibidos, nos damos cuenta de la cantidad de horas que pueden ser utilizadas en esta tarea por las personas que trabajan en la misma. Desde ordenanzas hasta ejecutivos, todos intervienen en la gestión de los documentos, ocupando muchas veces más tiempo del que sería deseable. Cuando la gestión de los documentos empieza a comer terreno a la propia actividad, hablamos de que las organizaciones se han "burocratizado". Cuando la gestión de los documentos se abandona hablamos de organizaciones ineficaces, donde nunca se encuentra nada y donde la información es un bien individual y no colectivo.

La sensación de que esto sucede cada vez más a menudo tiene sus bases en una realidad de la que todos participamos:

- La cantidad de información que se maneja en el mundo actual es cada vez mayor. Es lo que los especialistas en información llaman desde principios de siglo "el crecimiento exponencial de la información". En nuestras oficinas se traduce en el incremento de documentos que se reciben, algunos de los cuales son indispensables para poder seguir ejerciendo nuestra actividad.
- Tradicionalmente la información que se manejaba en las oficinas se encontraba en documentos en papel, pero actualmente el concepto de documento se ha ampliado a otros muchos soportes capaces de contener información: disquetes, ficheros informáticos, imágenes, audiovisuales, videos, fotografías, microfilms, etc., lo que supone una mayor dificultad en el control de los documentos.
- La extensión de la microinformática y la comunicación vía fax en los entornos ofimáticos ha supuesto en algunos casos la ruptura de sistemas tradicionales de gestión documental. No es poco frecuente encontrar documentos iguales

gestionados de distinta forma sólo porque primero se recibieron por fax y luego por correo; o secretarías que no son capaces de recuperar un documento almacenado en el disco duro del ordenador, si no han sacado una copia en papel para su archivo tradicional.

- En el momento actual existen grandes presiones competitivas, lo que unido a la orientación de servicio al cliente nos hace valorar cada vez más la rapidez de los tiempos de respuesta, y la gestión de los documentos o "papeleo" aparece muchas veces como un factor negativo.

Además, los especialistas en Documentación han hecho poco caso de esta realidad, estableciendo diferencias importantes entre los bibliotecarios y/o documentalistas encargados de gestionar la documentación llegada a las empresas desde el exterior y los archiveros encargados de gestionar la documentación generada desde dentro. Visto de esta forma, sólo las grandes empresas pueden permitirse el lujo de tener "centros de documentación" y algunas incluso "archivo centralizado", cuando la gestión de la información se plantea como una necesidad para cualquier tamaño de empresa.

El concepto Gestión Integrada de la Documentación no es nuevo, aunque creo que se ha tratado hasta ahora más de un concepto teórico que de una realización práctica.

Un modelo de Gestión Integrada de la Documentación se basa fundamentalmente en dos premisas:

- Forman parte del sistema tanto los documentos internos como externos en cualquier soporte, puesto que lo que se considera de importancia para la organización es la información que contienen.
- La información producida o recibida en el desarrollo una actividad dentro de una organización debe poder ser compartida (salvo los casos de necesaria confidencialidad) por todos los miembros de la misma que realicen la misma o similar actividad, con la finalidad de mejorar y evitar duplicidades en los procesos de trabajo.

y se desarrolla mediante un plan establecido-que contempla los siguientes puntos:

- La definición de los tipos documentales.
- Los sistemas de control y flujo de los documentos.
- Los sistemas de recuperación de la información mediante sistemas de indización y clasificación.
- Los criterios y sistema de expurgo de la documentación no válida.
- Las reglas para la denominación de los documentos, ficheros y archivos.
- Los niveles de confidencialidad de los documentos.

- La forma de almacenamiento de la documentación válida.
- Los manuales dónde se incluyen cada uno de los procedimientos para llevar a cabo los puntos anteriores.

2. TECNOLOGIAS EXISTENTES

En la actualidad, la puesta en marcha de Sistemas de Gestión Integrada de la Documentación (SGID) cuenta con un importante factor a su favor: la creciente importancia en el mercado de aplicaciones informáticas para la gestión de documentos. El desarrollo de las tecnologías de digitalización de imágenes y escaneado, de almacenamiento en archivo óptico, el desarrollo de los interfaces gráficos, etc., junto con el abaratamiento de los costes de los equipos informáticos han permitido que la aparición en el mercado de muchas soluciones ofimáticas para la gestión documental, que tienen como objetivo final la "oficina sin papeles".

Las aplicaciones informáticas de Sistemas de Gestión Integrada de la Documentación están compuestas por una gran variedad de tecnologías diferentes que integradas consiguen dar respuesta a cada uno de los casos concretos. La terminología utilizada para la promoción de los productos crea una cierta confusión, pues nos encontramos con muy distintas denominaciones para productos similares: Archivo electrónico de documentos, Gestión documental electrónica, Archivo óptico de documentos, Sistema de gestión de imágenes, etc. Los requisitos que debe cumplir un sistema son tan variados como tipos de organizaciones o empresas existen, por lo que una de las características que más se valora en estos sistemas actualmente es su flexibilidad y posibilidad de integración con otras aplicaciones.

Para intentar aclarar un poco el panorama describiré las principales funcionalidades que pueden abarcar estos sistemas:

- *Gestión de imágenes.* Se trata de la tecnología que permite el escaneado, captura y digitalización de la imagen de las páginas que componen un documento. Los escáner son las máquinas que permiten la captura de las imágenes. Funcionan mediante una fuente de luz que ilumina el documento, de forma que las partes negras de la página absorben la mayor parte de la luz, mientras que las blancas la reflejan. La luz reflejada por el documento es detectada por un dispositivo de células fotosensibles, que la convierten en carga eléctrica cuya magnitud corresponde a la intensidad de luz recibida. Estas cargas son convertidas en señales digitales, capturadas en un entramado de puntos. Cada uno de los puntos es denominado "pixel" y la imagen así recogida se dice que está bit-mapped (o representada en bits). De esta forma, la mayor resolución de una imagen depende de un mayor número de pixels (medida normalmente en pixels por pulgada). Para el escaneado en color el proceso es similar pero separando los componentes de los colores primarios y utilizando un sistema similar a la reproducción en color. En las imágenes en las que solo se recoge blanco y negro, un pixel ocupa un solo bit. Utilizando las escalas de grises, 1 pixel puede equivaler a 8 bits y en la

representación en color todavía más, dependiendo del grado de resolución.

Cada fichero de imágenes debe contener, además de la imagen representada en bits, información adicional que permita al programa que la use visualizarla y decodificarla. Esta información se almacena normalmente como una cabecera. Aldus y Microsoft han desarrollado un formato bastante extendido para el almacenamiento de imágenes, el formato TIFF (Tagged Image File Format), en el que la información relativa a las imágenes se almacena en 45 campos o etiquetas, que además quedan reflejados en un directorio.

Los requerimientos de memoria para el almacenamiento de imágenes son bastante importantes, por lo que utilizan las tecnologías de compresión/descompresión de imágenes utilizadas para las transmisiones por fax (grupo 3, grupo 4).

Para visualizar las imágenes escaneadas, el proceso se realiza de forma inversa, visualizando la imagen en la pantalla del ordenador. Lo ideal sería visualizar las imágenes con la misma resolución que han sido escaneadas, lo que en muchos casos supone la utilización de monitores de alta resolución (17 ó 19 pulgadas de 1.000 a 2.000 líneas). Además, en muchas aplicaciones se utilizan dispositivos de aceleración de imágenes para que la visualización en pantalla sea más rápida.

Desde la pantalla se pueden realizar gran cantidad de funciones en todas las aplicaciones, tales como aumentar o disminuir de tamaño o recortar o señalar una parte de la imagen.

De estas imágenes pueden reproducirse copias en papel mediante impresoras láser o plotters en el caso de planos o gráficos.

- *Tecnologías de almacenamiento óptico.* A pesar del abaratamiento espectacular de los costes del almacenamiento magnético (para el millón de caracteres, de \$2.500 en 1970 a \$11 en 1993) y de la todavía vigencia del microfilm como medio de almacenamiento masivo, el almacenamiento óptico se plantea cada vez más como la solución que acompaña a los sistemas de gestión documental.

El tipo de almacenamiento óptico más conocido y que se ha popularizado como medio de edición es el CD-ROM. Basados en la misma tecnología, los medios de almacenamiento óptico que se ofrecen usualmente con los sistemas de gestión documental son los discos WORM y los discos ópticos regrabables.

Los discos WORM (Write Once Read Memory) permiten a sus usuarios grabar sus datos y/o imágenes y luego recuperarlas. La grabación de los datos se hace produciendo cambios irreversibles en la superficie del disco, de forma que sólo puede grabarse una sola vez. No es necesario que todos los datos se graben de una vez, sino que se pueden ir añadiendo a las superficies no utilizadas del disco. Existen discos WORM de varios tamaños, siendo los comercializados más usualmente los de 5.25 y 12 pulgadas. Los discos de 5.25 tienen una capacidad de 325 Megabytes por cada lado y los de 12 alrededor de 1 Gigabyte, aunque los

desarrollos para poder leer ambos lados al mismo tiempo aumentan considerablemente la capacidad de almacenamiento. Existe también una versión WORM del CD-ROM, llamada CD-R, que tiene la ventaja de poder ser leída en los lectores estándar de CD-ROM, lo que abre grandes posibilidades en la distribución externa o interna de información.

Los discos ópticos regrabables emplean una forma magneto-óptica de grabación. Esto permite utilizarlos como los ya tradicionales discos magnéticos, pero con una capacidad de almacenamiento mucho mayor. Existen en versión de 5.25 y 3.5 pulgadas. Existen ya lectores que permiten la utilización de discos WORM y regrabables indistintamente, aunque su precio es todavía elevado.

Tanto los discos WORM como los regrabables son discos removibles, por lo que se han desarrollado dispositivos que permiten el cambio de un disco a otro sin intervención manual. Son los jukeboxes, que permiten mantener de 10 a 10.000 discos diferentes que pueden ser leídos desde 1 a 100 bocas diferentes. Con estos sistemas se permite tanto consultar varios discos al mismo tiempo, como cambiar automáticamente el disco que se necesita.

- *Tecnologías de reconocimiento.* Los documentos escaneados son guardados y recogidos como simples imágenes. Cuando queremos recuperar el texto de los documentos para poderlo tratar posteriormente como trataríamos un texto realizado con nuestro procesador de textos, debemos utilizar las tecnologías de reconocimiento de caracteres. Estos programas actúan reconociendo cada uno de los caracteres del texto del documento escaneado y convirtiéndolo en un texto ASCII.

Las tecnologías de reconocimiento incluyen la lectura de códigos de barra, los OCR (reconocimiento óptico de caracteres) y los ICR (reconocimiento inteligente de caracteres). Los programas OCR son capaces de reconocer las letras impresas de fuentes estándar, mediante la comparación con una serie de imágenes de caracteres que han sido pregrabadas. Los ICR no funcionan buscando la imagen pregrabada sino las reglas que pertenecen a cada uno de los caracteres, de forma que son capaces de distinguir las letras de diversas fuentes y estilos de impresión. Siguiendo en esta línea de desarrollo se investiga en el reconocimiento de la escritura manual.

El reconocimiento de caracteres es enormemente útil cuando tratamos documentación normalizada, como puede ser el caso de digitalización de formularios. En este caso, la información que sirve como índice de las imágenes puede ser extraída de un lugar concreto dentro del formulario (casillas con número de expediente, por ejemplo). Este proceso puede hacerse mediante OCR en batch y sin necesidad de intervención humana.

Para los dibujos técnicos y de diseño, las imágenes deben ser convertidas al formato vectorial o de vectores que utilizan las aplicaciones CAD (Diseño asistido por ordenador). Estos programas de vectorización son ofrecidos en las

aplicaciones especialmente diseñadas para el tratamiento de este tipo de imágenes.

- *Software de gestión de documentos.* Para la recuperación de las imágenes de los documentos almacenadas es además necesaria la creación de índices. Estos índices están compuestos por las distintas claves de búsqueda por las que se quiere localizar el documento, que quedan unidas a la/s imagen/es mediante una clave de enlace. Por lo tanto, es necesario que las aplicaciones de gestión de documentos incluyan programas de bases de datos que nos permitan la construcción de índices y, por lo tanto, la recuperación de las imágenes de los documentos.

Dependiendo del tipo de información que necesitemos recuperar y del tipo de documentación a tratar, podemos decantarnos por modelos de bases de datos relacionales o de bases de datos documentales. Las primeras se ajustan mejor a la información estructurada, ya que requiere la grabación de un registro (cuyos campos han sido diseñados previamente) para cada documento que se incluya en el sistema informático.

Las bases de datos documentales se adaptan mejor cuando lo que necesitamos es la recuperación por palabras clave del texto, que pueden ser recuperadas de una parte o de todo el texto completo. Este tipo de recuperación incluye la posibilidad de utilización de los operadores booleanos y de la utilización de una lista de palabras vacías. Igualmente puede ser de mucha utilidad la posibilidad de soportar un tesoro, con el que se contrastan los términos de indización.

El tipo de software utilizado en estas aplicaciones no difiere del tipo de software tradicionalmente usado en aplicaciones para el tratamiento de bases de datos referenciales. Lo que sí que es importante en este tipo de aplicación es la capacidad del software de tratar de la misma forma documentos en distintos soportes. Es decir, no solamente tener la capacidad de tratar imágenes captadas de documentos en papel sino de ser capaces de integrar estas con documentos electrónicos producidos por distintas aplicaciones informáticas utilizadas en las oficinas: procesadores de texto, hojas de cálculo, etc.

El software de base de datos de desarrolla en estos momentos hacia aplicaciones de bases de datos que reúnan las dos posibilidad de recuperación, a las que además se pueden añadir funcionalidades del software hipertexto e hipermedia.

Una de las últimas soluciones ofrecidas es la utilización de bases de datos orientadas al objeto para la gestión de los contenidos de los documento. Sin embargo, en cómo se debe desarrollarse esta tecnología no todos los expertos se ponen de acuerdo. Por un lado se piensa que deben desarrollarse nuevos lenguajes de recuperación para este tipo de bases de datos y por otro las bases de datos tradicionales se desarrollan hacia la inclusión de conceptos de la orientación a objetos.

- *Software de gestión de los flujos de trabajo (workflow)*. El verdadero beneficio de la instalación de este tipo de aplicaciones se encuentra cuando sirve para ayudar al personal de una organización a trabajar en conjunto, a compartir información, a hacer más rápidos los procesos y a mejorar su productividad. Para ello se ha desarrollado software específico que, basado en la definición previa de una serie de tareas y procedimientos, asegura que las tareas son repartidas entre las personas correctas, en la secuencia adecuada y que son acabadas sin sufrir demoras innecesarias. De esta forma, cuando un documento entra en la organización es escaneado e identificado, y a partir de aquí el sistema sabe quién debe ver ese documento y cómo debe circular dentro de la organización. Lo envía a la cola de trabajo de la primera persona que lo debe ver, y si en un cierto período de tiempo no se ha trabajado con él, el sistema le recuerda que debe hacerlo y en qué fecha debe estar terminado. Cuando la primera persona finaliza su trabajo es el sistema el que se encarga de hacerlo llegar a la siguiente en la cadena de trabajo.

Estos sistemas permiten además llevar un importante control de los procesos de trabajo y medir la productividad de los individuos, departamentos u organizaciones.

- *Software para trabajo en grupo (groupware)*. Se relaciona estrechamente con el anterior, incluyendo entre otras prestaciones las siguientes:
 - . Correo electrónico
 - . Calendarios compartidos
 - . Videoconferencias
 - . Edición colaborativa de documentos en tiempo real
 - . Gestión compartida de documentos

3. RAZONES PARA LA INSTALACION DE SISTEMAS INTEGRADOS DE GESTION DOCUMENTAL

Los SGID, que como hemos visto funcionan como integradores de diferentes tecnologías, son desarrollos cuya finalidad es ayudar a las empresas y organizaciones a conseguir sus objetivos. No constituyen una finalidad en sí mismos, sino que contribuyen a mejorar la eficacia y efectividad de cada campo de actividad. Por lo tanto, su instalación debe ser considerada dentro del contexto de la estrategia o sistema de información de cada empresa, y este último debe estar totalmente integrado dentro de la planificación general.

A pesar de que los SGID son lo suficientemente nuevos para no existir experiencia

acumulada que nos permita analizar en profundidad los resultados de su instalación, podemos observar que las organizaciones que han dado este paso (sobre todo en USA) se han basado en un deseo de ganar competitividad. Los primeros resultados indican que se ha conseguido.

Las razones que se pueden esgrimir para la instalación de un SGID deben ser estudiadas en cada caso. Su instalación todavía no es barata y puede suponer un cambio importante en los procesos de trabajo y en la forma de trabajar, por lo que es absolutamente necesario realizar un proceso previo de planificación, en el que se definan y cuantifiquen los objetivos perseguidos con su implantación.

Presentando un enfoque general, las razones pueden dividirse en tres grandes bloques:

- *RAZONES ESTRATEGICAS*, que incluyen consideraciones sobre la marcha general de la empresa y afectan a la mayoría de sus áreas:

- . Hacer más rápidos los procesos de trabajo en los que interviene el manejo de documentación
- . Asegurarse de que la documentación no es el factor que retrasa los procesos de trabajo
- . Flexibilizar la organización física de la documentación y las personas
- . Incrementar la posibilidad de crecimiento, que está limitada por la capacidad de procesar más documentación en papel
- . Mejorar el servicio al cliente
- . Mejorar la satisfacción en el trabajo del personal
- . Mejorar la seguridad de los documentos contra incendios e inundaciones
- . Establecer procesos de auditoría automáticos
- . Utilizar un mejor sistema de archivo

- *RAZONES FINANCIERAS*, que inciden en la reducción de costes y aumento de los ingresos:

- . Incrementar la productividad, reduciendo el tiempo necesario para realizar determinados procesos: reduciendo los recursos necesarios, produciendo más con los mismos o aumentando la calidad con los mismos recursos.
- . Ahorrar espacio físico, eliminando todos los espacios dedicados al archivo de documentos en papel, así como los espacios de fotocopiadoras o

máquinas de microfilm.

- . Incrementar el volumen de negocio al poder poner los productos antes en el mercado, o al mejorar la calidad o al aumentar la capacidad de creación de nuevos productos.
- . Reducir los gastos corrientes, tales como los gastos de fotocopias, de devolución de llamadas, de microfilmación o de mensajeros.

- *RAZONES TECNICAS*, que presentan la instalación de un SGID como la solución a una serie de problemas existentes dentro de la organización:

- . Deficiente gestión de la documentación en papel, que supone dificultades y demoras en la recuperación de la información, documentos de gran valor que no son explotados en su totalidad, pérdida de documentos, etc.
- . Difícil integración de la información en distintos soportes.
- . Lenta comunicación de la información.
- . Duplicidad en el tratamiento de la información que es procesada en sistemas distintos.
- . Retrasos en el procesamiento de la información.

4. LOS ESPECIALISTAS EN INFORMACION Y LOS SGID

Como ya he dicho anteriormente, cualquier proceso de implementación de un SGID necesita de una planificación importante. También como es lógico se requiere un seguimiento, evaluación y puesta al día del sistema una vez que se ha implantado.

Este proceso de planificación requiere un esfuerzo de colaboración no demasiado frecuente entre los directivos y los técnicos, puesto que el simple hecho de reestructurar los sistemas de proceso de la información pueden hacer revisar toda la filosofía de la empresa. Entre los técnicos involucrados en estos procesos deben encontrarse tanto especialistas informáticos como especialistas en organización y especialistas en documentación e información. Esta simbiosis pocas veces se produce en nuestro país, aunque sea obviamente beneficiosa.

Los especialistas informáticos utilizarán sus conocimientos para diseñar las soluciones técnicas más adecuadas en cuanto a las máquinas, las redes de comunicaciones y la adecuación de los programas necesarios. En muchos casos también se tendrán que realizar desarrollos a medida para casos particulares. Los especialistas en información aportarán a estos sistemas sus conocimientos sobre la recuperación de la documentación, la tipología documental, la difusión de la información y la selección de

fuentes documentales. Y los especialistas en organización desarrollarán sus conocimientos en el desarrollo de nuevos procedimientos de trabajo. Todos ellos actuando en consonancia con los objetivos que han sido desarrollados por la dirección.

5. BIBLIOGRAFIA

* Document Management Yearbook 1994: A guide to imaging and document management products and services / compiled by Roger N. Broadhurst. Hertfordshire: CIMTECH, UKAIM, 1993

* Green, William B. Introduction to electronic document management systems. Boston: Academic Press: 1993

* Haapaniemi, Peter. The power of imaging. Electronic Library, v.8, n.6, December 1990, p.401-407

* Wiggins, Bob. Document Image Processing: An overview. Document image automation, v.12, n.3, Fall 1992, p. 3-9 ;(part 1); v12, n.4 Winter 1992, p. 12-20 (part 2)