



Prototipo de captura automatizada de documentos y gestion de contenidos ECM.

PCCM

04/10/12

Índice de contenido

1. INTRODUCCION-OBESIDAD DOCUMENTAL.....	3
1.1. Análisis del ROI.....	4
1.2. Conclusiones sobre el ROI.....	6
2. PCCM (Prototype Content and Capture Management).....	7
.....	7
2.1. Características del PCCM.....	9
2.1.1. ¿Que prototipos se pueden solicitar?.....	10
2.1.2. ¿Qué ventajas nos aporta el PCCM?.....	11
3. Propuesta Económica.....	12
3.1.Observaciones:.....	12

1. INTRODUCCIÓN-OBESIDAD DOCUMENTAL

Recientemente, George Parapadakis ha acuñado el término de **Obesidad documental y de contenidos**.

¿Que significa realmente este término?

Es una excelente analogía sobre la condición en la que se encuentran las organizaciones. Según un estudio realizado por IDC ([Http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf](http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf)) para el 2020 el volumen de información almacenada se incrementará por 50 veces sobre la información existente en 2011.

En la actualidad existe un exceso y redundancia de la información, el mismo contenido puede estar almacenado n veces en el ecosistema de la organización. Esta circunstancia evidentemente, provoca un efecto adverso en la eficiencia y eficacia del negocio porque se generan unos altos costes y riesgos para la organización.

- Tiempo dedicado a archivar los contenidos.
- Tiempo dedicado a la localización de los contenidos deseados.
- Costes de almacenamiento físico de los contenidos.
- Los costes asociados a la distribución de los contenidos.
- Los costes asociados a la introducción de miles de datos de forma manual a los sistemas de información, errores de procesamiento.
- Los costes relacionados con la redundancia de los contenidos.
- Backups de los contenidos.

Esta imparable acumulación de contenidos impide que las organizaciones se muevan con agilidad, se vuelven cada vez mas pesadas. Por lo tanto la solución no puede venir en ningún caso de la mano de una reducción de los contenidos, al contrario, cada vez van a crecer más, emails, documentos, webs, proyectos, vídeos...

¿Que alternativas tienen las organizaciones?

En los últimos años, las organizaciones han realizado grandes esfuerzos en controlar los procesos de gestión relacionados con las ventas, compras, financiero, relación con los clientes, etc... han aprendido a introducir datos y en el mejor de los casos a analizar los indicadores de negocio a través de herramientas de B.I.

A pesar de que se han hecho grandes esfuerzos en TICs, ciertas cosas nunca cambian, se guardan los documentos en AZ y nunca asociados al dato digital, seguimos procesando los documentos de forma manual, seguimos teniendo dificultades para buscar el documento que necesitamos..

La solución no viene de una tecnología sustitutiva sino de una complementaria que nos va a permitir mantener toda la información en constante flujo, que nos va a permitir recuperarla de forma rápida y evitar la redundancia de los contenidos.

La solución viene a través de lo que denominamos Gestión de Contenidos Empresariales o ECM.

¿Por qué interesa a las organizaciones una solución de gestión documental y contenidos con captura de documentos automatizado?

Porque se obtiene un ROI elevado. Ahora bien, aunque Binovo esta convencido de que este tipo de soluciones proporcionan a las organizaciones unos índices de eficiencia considerables, nos queda demostrar que efectivamente, la inversión realizada cubrirá las expectativas de las organizaciones en cuanto una mejora significativa en la gestión de los contenidos empresariales y automatización en la captura de contenidos.

1.1. Análisis del ROI

El ROI es el ratio resultante de $(\text{Beneficio obtenido} - \text{Coste}) / (\text{Coste})$.

¿Qué costes y beneficios debe analizar una organización?

Costes:

1. Analizar los costes derivados del trabajo. Es decir, el objetivo es poder calcular cual es coste anual de gestión de documentos. Tenemos que tener en cuenta estos factores:
 - Costes salariales, ¿Cuanto me cuesta el trabajador por minuto?
 - Gestión de documentos. Hay que calcular el tiempo medio (Minutos) dedicado a la gestión de un documento.
 - Recepción de documentos. Identificación.
 - Tratamiento, introducción en sistemas o bases de datos.
 - Gestiones internas y externas relacionadas con el documento.
 - Archivado del documento.
2. Analizar los costes por los errores cometidos en el tratamiento de los documentos. En ocasiones las organizaciones no valoran los costes de gestión (Coste/ hora/ empleado) que se genera cuando la información introducida en los sistemas de información (ERP, CRM...) no es la correcta o es errónea.

-
3. Analizar los costes derivados del almacenamiento físico y digital. El objetivo es llegar a calcular el coste anual de almacenamiento.
 - Número de armarios archivadores.
 - M2 destinados al almacenamiento de los contenidos físicos.
 - Servidores y sistemas de almacenamiento digitales para archivo de backups y contenidos redundantes.
 4. Analizar los costes de búsqueda y distribución de los documentos. El objetivo es calcular de forma anualizada el tiempo que los operarios tardan en buscar información y distribuirla.
 5. Analizar el coste que afecta a la pérdida de los documentos.
 6. Analizar otros costes que pueden ser sectoriales o específicos de una organización.

Ahora que ya podemos calcular los costes, ¿Qué beneficio pueden obtener las organizaciones en el caso de que incorporen soluciones en la captura y gestión de contenidos que les permita ser mas eficientes?

Beneficios

Los beneficios atribuibles a las soluciones de captura y gestión de contenidos se pueden resumir en:

1. Se reduce el coste de transporte de los documentos en papel. Proceso de captura de documentos de arquitectura distribuida.
2. Reducción de los costes de proceso. La captura automatizada identifica los documentos, clasifica, extrae la información mediante tecnología de OCR, ICR, OMR y códigos de barra (1D y 2D), se permite la validación (para evitar errores) en función de miles de reglas y se exportan los contenidos junto a los metadatos para su explotación. Mínima intervención humana.
3. Se reducen los costes por la mejora en la capacidad de búsqueda y recuperación de los documentos. Según un estudio de IDC, los trabajadores de oficina suelen consumir entre un 15% y 35% de su jornada buscando documentos.
4. Costes relacionados con la seguridad y accesibilidad a los documentos y la mantenibilidad legal. Los documentos electrónicos perduran en el tiempo y garantizan su accesibilidad ante contenidos con límites de almacenamiento legal.
5. Costes relacionados con la distribución de los contenidos. Los ECM facilitan el trabajo colaborativo y la creación de flujos de trabajo que mejoran la eficiencia de los procesos documentales. Evitamos la información redundante porque se trabaja de forma colaborativa sobre el contenido.

-
6. Reducción de los costes de almacenamiento a medida que los contenidos se vayan digitalizando teniendo en cuenta las prácticas y legalidad vigente sobre la conservación de los documentos y firma electrónica.

Aunque estamos cerca de poder calcular el ROI de un proyecto de digitalización, captura y gestión de contenidos, es necesario saber cuál es la inversión que vamos a llevar a cabo.

Inversión

1. Los costes de licenciamiento y servicios necesarios para implementar la solución en la organización.
2. Costes de inversión en infraestructura de almacenamiento, escaners, hosting o la arquitectura idónea para que funcione la solución.
3. Costes asociados a la resistencia al cambio.

1.2. Conclusiones sobre el ROI

La clave está en saber cuando recuperamos la inversión realizada, teniendo en cuenta los beneficios que hemos obtenido por la reducción de costes. Ganamos en eficiencia y eficacia, eso es por lo menos lo que nos quieren transmitir los expertos en la gestión de contenidos.

Lo que se evidencia es la dificultad de calcular los beneficios que se atribuyen a la captura y gestión de los contenidos hasta que de alguna manera lo podamos demostrar.

Binovo ofrece a este respecto una propuesta que permite valorar mediante un prototipo, los beneficios que se atribuyen a estas soluciones.

El PCCM (Prototype Content and Capture Management) es el camino para que las organizaciones puedan medir de forma real, los beneficios y bondades de la captura y gestión de contenidos, utilizando un extracto documental significativo que pueda extrapolarse al resto de las actividades sin apenas coste.

2. PCCM (Prototype Content and Capture Management)

Se trata de una herramienta que Binovo pone en manos de las organizaciones para utilizar durante un periodo de tiempo las soluciones de captura y gestión de contenidos focalizados principalmente a un tipo de documentos como son las facturas de proveedores o albaranes de proveedores. El prototipo estará abierto a otros tipos de contenidos, siempre que la organización considere que puede ser mas representativo y pueda valorar su impacto en la organización.

La solución estará formada por:

DATACAP TASKMASTER

Datacap es una plataforma de captura de IBM que permite automatizar la entrada de datos a los gestores de contenidos y sistemas de información.



-Escanear o capturar contenidos ya digitalizados: el sistema identifica el contenido y realiza las tareas necesarias de transformación para que la tecnología de extracción sea lo más eficaz posible.

-Reconocimiento: mediante las tecnologías de reconocimiento, OCR , ICR, OMR, Barcode (1D y 2D), Datacap extrae la información de los documentos de forma automática.

-Verificación: mediante un potente engine de reglas, los usuarios pueden validar los datos extraídos. Ejemplo: que se pueda leer un DNI, pero mediante la regla de cálculo del DNI poder indicar si es válido o no, o no poder validar una factura hasta que la Base + IVA no sea igual al total.

-Exportación: una vez validados los datos el sistema puede exportar los datos a una BBDD mediante ODBC, puede crear un XML o puede crear los contenidos con los metadatos en un sistema ECM.

ECM ALFRESCO

Alfresco es una plataforma ECM (Gestión de Contenidos Empresariales) que permite una gestión eficiente y efectiva de los contenidos de una organización.

¿Cuál es su función dentro de la solución?

La función de Alfresco es almacenar y gestionar los contenidos que se generan en la organización.

- **Búsqueda:**
 - **Indexación:** mediante su potente indexador, los contenidos que se generan y se almacenan se indexan, haciendo posible una búsqueda mucho más eficaz. Todos los contenidos tipo texto se indexan sin ningún proceso previo de OCR (odt, doc,xls, etc.); en contra, los contenidos tipo imagen necesitan de un proceso previo que permita convertirlo en texto. Ese proceso lo hace Datacap.
 - **Metadatos y etiquetas:** es otra de las características que permiten identificar los contenidos para facilitar su búsqueda o realizar acciones automatizadas a través de flujos de trabajo. La herramienta Datacap es capaz de generar estos metadatos de forma automatizada.
- **Sistema de reglas y notificaciones:** *Alfresco* proporciona un mecanismo de reglas para las carpetas. Este mecanismo permite desencadenar automáticamente ciertas acciones sobre documentos que se añaden, eliminan o modifican en una carpeta si estos cumplen ciertas condicionantes.
- **Múltiples mecanismos de acceso:** *Alfresco* permite el acceso a la documentación por varias vías de acceso, FTP o CIFs, IMAP y SMTP.
- **Integración *Active Directory* y *Single Sign On*.**
- **Gestión de permiso y control de acceso.**

En el siguiente enlace podemos ver un vídeo que detalla perfectamente lo que pretendemos conseguir con la solución.

[Http://www.binovo.es/ftp/jbastarrika/Datacap/ScreenCastTake01.avi](http://www.binovo.es/ftp/jbastarrika/Datacap/ScreenCastTake01.avi)

2.1. Características del PCCM

La organización que solicite un PCCM, obtendrá los siguientes servicios.

1. Visita de un técnico especialista en captura de contenidos y gestión de contenidos para realizar las siguientes tareas:
 - Selección de un muestreo de albaranes o facturas de proveedores.
 - Instalación de Datacap en un ordenador de características normales actuales, mínimo de 2 Gb de memoria Ram. Es necesario tener un escaner.
 - Instalación del conector que permitirá subir los contenidos al ECM Alfresco, que estará ubicado en un servidor remoto de Binovo.
 - Parametrizar Datacap para que exporte los contenidos capturados en un fichero excel. Esta funcionalidad será equivalente a la integración de Datacap con el sistema de gestión de la empresa.
 - Pequeña formación de como utilizar Datacap.
2. La duración del PCCM será de un mes prorrogable a 2 meses.
3. Una vez finalice el periodo de pruebas, Binovo se pondrá en contacto con la organización para valorar y sacar conclusiones de las pruebas realizadas.
4. Si la organización concluye una vez que ha extrapolado el muestreo al resto de la organización que la solución le convence, Binovo realizará una propuesta definitiva de la solución que desea la organización.
5. En la propuesta final se tendrán en cuenta los datos que Binovo ha podido ir recogiendo en el periodo de prueba.
 - Se ofertara la arquitectura adecuada para cubrir sus necesidades.
 - Integración con sus sistemas de información, incluso con su gestor documental en el caso de que no sea Alfresco.
 - El coste del PCCM se descontará de la oferta final.
 - La organización se podrá beneficiar del plan de financiación de proyectos que ofrece IBM, siempre sujeto a valoración de riesgo, que supone poder financiar la totalidad del proyecto a un año sin intereses o hasta 4 años con intereses.

2.1.1. ¿Que prototipos se pueden solicitar?

En principio se ha definido un prototipo estándar para un tipo de contenido, facturas y albaranes de proveedores.

La decisión es bien sencilla, los directivos de las empresas, perciben que las organizaciones dedican mucho tiempo a introducir las facturas de proveedores, gestionar los pagos y errores provocados en la gestión, con lo que los costes son considerables. Los usuarios mediante el ejemplo de las facturas podrán hacerse una idea mas realista de las posibilidades de automatización que nos proporcionan las herramientas de captura y de gestión de los contenidos.

No obstante, aquellas organizaciones que deseen prototipar otro tipo de documentos, formularios, contratos, seguros, extractos bancarios etc... lo podrán solicitar.

Prototipo de factura de proveedores:

1. El usuario escaneará las facturas de proveedores y los guardará en un directorio determinado.
2. Datacap procesará los lotes de facturas, identificando, clasificando y aplicándole las tecnologías de extracción de datos OCR.
3. Si Datacap no reconoce el formato de la factura de proveedores, el sistema así lo indica y el usuario mediante clicks le indicará que campos debe leer para que aprenda que debe hacer la próxima vez que detecte ese formato.
4. Una vez se hagan las tareas de extracción de los datos, el sistema mostrará la ventana de validación de los datos. Es el proceso por el cual el usuario dará por buenos los datos que Datacap ha extraído. En este proceso de validación existen reglas de validación que han sido programadas con anterioridad y que ayudarán al usuario a que no se almacenen facturas con datos erróneos: validación de CIFs, de cantidades e impuestos etc...
5. Una vez el usuario valide los datos, se exportará automáticamente :
 - Generará un Excel con los datos de las facturas que se han extraído. Al igual que hace esta tarea se pretende demostrar que se puede integrar con ERPs, por ejemplo, creando las factura de proveedor en el módulo de compras de su ERP.
 - La imagen de la factura junto a sus metadatos (CIF, número de factura, fecha, base imponible, impuestos y total) se exportarán a Alfresco, que estará ubicado en los servidores de Binovo.
6. El usuario podrá acceder a Alfresco para poder buscar las facturas por su contenido y por los metadatos que el proceso de captura ha extraído.

2.1.2. ¿Qué ventajas nos aporta el PCCM?

1. La organización puede probar con sus documentos, el nivel de automatización que se puede incorporar en la gestión de sus contenidos. Todo esto por un coste muy reducido.
2. Los costes de adaptación a la herramienta son mínimos.
3. Se elimina parte de la incertidumbre que se genera cuando una organización decide incorporar este tipo de tecnología porque no incurrimos a grandes inversiones y existe un punto de retorno.
4. Las pruebas se pueden utilizar para extrapolar los resultados al resto de los departamentos de la organización y medir los resultados de hacer las tareas en la forma tradicional o a través de la nueva solución.
5. Proporciona el mecanismo para calcular un ROI aproximado que ayude a tomar una decisión sobre la inversión.