Productividad10

www.clase10.com

E-BOOK

Gestión del ciclo de vida de los datos

Creación y captura; transmisión, almacenamiento y seguridad; gestión y trabajo colaborativo; análisis y explotación de los datos aplicada al negocio

C10 | Del Dato a la Ventaja Competitiva

Contenidos

Introducción	3
Conceptos previos: tipos de datos	4
Fases de la gestión de vida de los datos	6
Caso práctico de aplicación de la Gestión del Ciclo de Vida de los Datos	13

Introducción

La aplicación de tecnología a los procesos productivos y empresariales permite a las empresas **ser más competitivas**, poder transformar la forma en que fabrican o prestan su servicio y establecer relaciones a largo plazo con sus consumidores en un escenario en el que las Tecnologías de la Información (TICs) cada día tienen un mayor peso.

Aplicar de forma inteligente e intensiva las Tecnologías de la Información a la gestión empresarial hoy en día, ya está ayudando a las empresas a tener **procesos de fabricación más ágiles y con un mayor nivel de productividad**, a reducir la dependencia de la constante reducción de costes para poder competir en mercados cada vez más globalizados y dinámicos; y a aportar valor añadido en los procesos productivos.

Ahora bien, el uso de la tecnología lleva aparejada la gestión de un volumen creciente de información que se genera tanto en las operaciones diarias de cada

compañía como en la interacción con sus clientes.

Asumir, tratar y procesar la información creciente supone todo un reto para las empresas y de una correcta gestión de los datos, depende el hecho de **transformar esa información en un activo clave para la organización y la toma de decisiones**.

La mayoría de los profesionales que gestionan datos pueden reconocer que existe un **ciclo de los datos**. Y un ciclo de vida comprende diferentes fases, cada una de ellas con sus propias características.

En este e-book vamos a desgranar cuáles son esas fases que implica la gestión del ciclo de vida de los datos.

“De una correcta gestión de los datos depende transformar la información en un activo clave para la organización y toma de decisiones”

Conceptos previos: tipos de datos

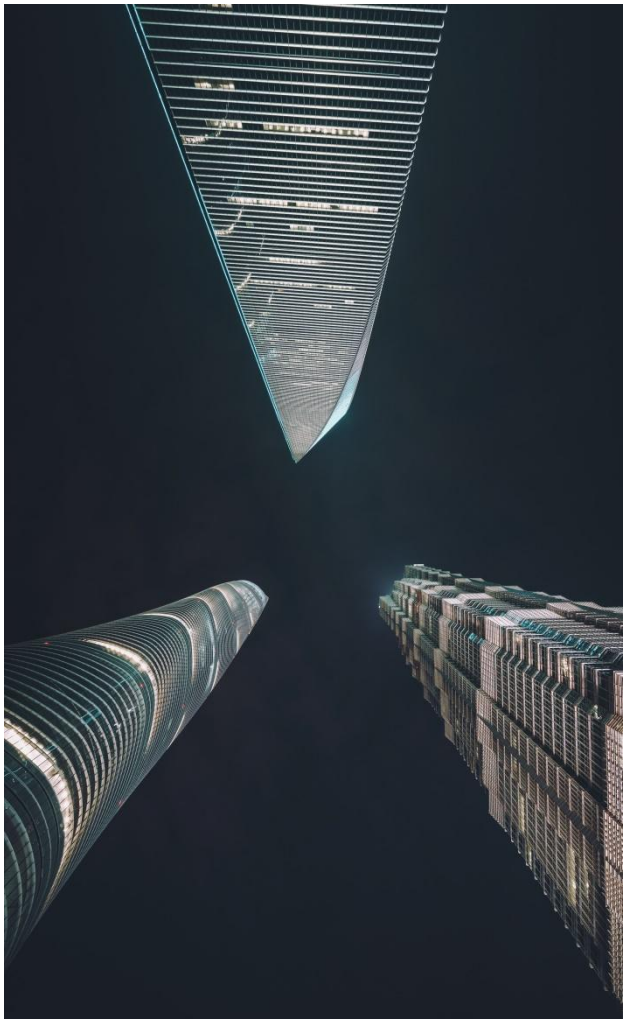
Como vamos a ver en la siguiente sección, la gestión del ciclo de vida de los datos requiere de todo un proceso que abarca desde su generación o captura hasta su almacenamiento y posterior explotación. Por tanto, en este sentido existen varios planteamientos de gestión o gobernanza de la información y algunos conceptos previos a tener en cuenta.

Así, los datos se pueden clasificar de acuerdo a dos tipologías diferentes:

- **Datos estructurados**. Engloba a toda la información y datos que se han creado u obtenido y que ha sido almacenada en tablas de bases de datos relacionales.

- **Datos no estructurados**. Estos hacen referencia a la información electrónica y datos que se han creado u obtenido y que no ha sido almacenada en las citadas tablas de bases de datos relacionales, incluyéndose todo tipo de información relacionada con e-mails, documentos de naturaleza ofimática, archivos PDF, documentos gráficos y visuales, etc.

> “La gestión del ciclo de vida de los datos requiere de todo un proceso que abarca desde su generación o captura hasta su almacenamiento y posterior explotación”



A la hora de gestionar los datos, uno de los planteamientos usados es conocido como *Data Lifecycle Management - DLM* (En castellano, Gestión del ciclo de vida de los datos). El DLM pone el énfasis en las diversas características que componen todo el contexto de los datos desde un enfoque empresarial y técnico: el diseño de la arquitectura, el desarrollo y tipo de la base de datos, la gestión de las diversas fases y procesos por los que un determinado dato tiene que pasar, las diversas medidas de seguridad y de qué forma es almacenado, etc.

Otra forma de gobernanza es la llamada **Gestión Documental** (Enterprise Content Management – ECM), modelo en el que los datos tienen principalmente formatos documentales y de ahí su asociación con la Gestión Documental.

La **tendencia actual** es que el mundo empresarial está demandando una forma de **gobernanza de la información que aúne ambas formas** y tipologías de datos bajo un mismo enfoque común que permita una integración y mayor facilidad y gestión empresarial.

Fases de la Gestión del Ciclo de los Datos

1. Creación y Captura

La primera fase a abordar en la gestión del ciclo de vida de los datos es su creación. El cómo se crean, reciben o capturan datos que entran a formar parte de toda la información del negocio. Esta primera fase se conoce como **creación y captura de datos**.

Los datos pueden crearse bien **a partir de operaciones que realizan las personas** -como puedan ser documentos de oficina, formularios, contenidos multimedia, etc.- o bien, **a través de los sistemas o las aplicaciones informáticas** como ERPs, los sistemas de facturación electrónica (E-billing), aplicaciones financieras, dispositivos (recepción de señales, RFID, IoT), etc.

En esta fase, se debe dar el siguiente tratamiento a los datos:

- **Agregación**: es el proceso de combinar la entrada de datos procedentes de diferentes herramientas de creación y otros sistemas.

- **Categorización**: organización de los datos, documentos, páginas web y otros contenidos en agrupaciones lógicas, basándose en el contenido y en la información que contienen.

- **Indexación**: identificación de atributos específicos de un documento o registro de base de datos de cara a facilitar la recuperación.

Para capturar y crear datos, se hace uso de distintas **tecnologías que permitirán crear o capturar datos y guardarlos en un repositorio para su reutilización, distribución, almacenamiento, análisis y explotación.**

- **OCR** (*Optical Character Recognition*): el reconocimiento óptico de caracteres no es más que una tecnología que permite la digitalización de textos, identificando caracteres de forma automática o unas imágenes símbolo que pertenecen a un alfabeto y transformándolos con forma de datos para su almacenamiento y uso.

- **ICR** (*Intelligent Character Recognition*): es una forma avanzada de OCR que puede incluir funcionalidades que permitan el aprendizaje de fuentes durante el procesamiento, el uso del contexto para mejorar el acierto en el reconocimiento o incluso el reconocimiento de escritura a mano.

HCR (*Handprint Character Recognition*): tecnología diseñada para convertir en código imágenes de caracteres escritos a mano.

- **OMR** (*Optical Mark Recognition*): es una tecnología que detecta la presencia o ausencia de marcas en áreas definidas. Se utiliza para pruebas estandarizadas, procesamiento de cuestionarios, etc.

- **Document Imaging:** es un proceso de captura, almacenamiento y recuperación de documentos independiente de su formato original, usando el escaneado, OCR, ICR, etc.

- **E-Forms/Web Forms**: ocurre cuando el diseño, gestión y procesamiento de los formatos se da completamente en un entorno digital o electrónico.

- **Forms processing:** funcionalidad de un software que permite procesar formatos con información y extraer datos para almacenarlos en una base de datos. El reconocimiento inteligente de documentos, identifica automáticamente los tipos de documentos desde el diseño y la estructura del documento.

- **COLD/ERM:** esta tecnología almacena e indexa la salida de datos del equipo (informes principalmente) en discos magnéticos o en discos ópticos. Esta información puede ser recuperada, vista, impresa o distribuida a Internet. Se utiliza a menudo para aplicaciones de facturación online.

Durante la primera fase, se crean datos estructurados o no estructurados.

Puede ser en forma .de un documento de MS Office 365, de un archivo en PDF, correo electrónico, información en una base de datos o incluso imágenes.

Por ello, en esta fase es importante definir a nivel tecnológico:

- Qué datos se requiere capturar o crear (datos, metadatos, documentos, etc.).
- Cómo se van a crear estos datos (formatos, dispositivos, etc.).
- Cómo se van a organizar los datos (organización de datos consistente).
- Cómo se van a almacenar los datos (sistemas de ficheros, repositorios, bases de datos, data warehouse).

2. Transmisión, almacenamiento y seguridad

En la segunda fase del ciclo de vida de los datos, se debe dar respuesta a preguntas como:

- ¿Dónde vamos a poner la información?
- ¿Cómo vamos a encontrar la información después?
- ¿Cómo se van a mover o transmitir los datos entre los sistemas corporativos o con sistemas externos?
- ¿Cómo vamos a garantizar la seguridad de la información?

A la hora de almacenar la información, se debe definir **cuál va a ser el sistema de ficheros**, es decir, cómo se van a nombrar los archivos y el esquema lógico a seguir para el almacenamiento y la recuperación. La estructura de colocación de los archivos más común es la jerárquica (árbol).

Los datos y los contenidos se pueden almacenar haciendo uso de distintos sistemas:

- **Bases de datos**: colecciones electrónicas de registros almacenados en un archivo central y accesible para muchos usuarios y muchas aplicaciones. Los datos de los registros o archivos tienen relaciones con otros registros o archivos.

- **Sistemas de gestión de contenidos**: sistemas software que tienen la capacidad de gestionar y rastrear la ubicación y las relaciones entre el contenido dentro de un repositorio.

- **Repositorio**: parte de un sistema de gestión de contenidos que incorpora funcionalidad específica para controlar la entrada y salida de información, control de versiones y búsqueda de atributos definidos, entre otros.

- **Almacén de datos**: repositorio central para todos o la mayoría de los datos estructurados de una organización.

En cuanto al “lugar”. donde almacenar la información, hasta hace poco más de una década era habitual el uso de dispositivos de tipo CD-ROM/DVD, almacenamiento magnético, redes de área de almacenamiento (SAN), almacenamiento conectado en red (NAS) y, en general, servidores locales. Sin embargo, especialmente en el último lustro, ha tenido un gran crecimiento el almacenamiento de información en la conocida “nube”.

El **almacenamiento de datos en la nube** es un modelo basado en redes de computadoras donde los datos se alojan en espacios de almacenamiento virtualizados, aportados por empresas que operan enormes centros de procesamiento de datos y se accede a los servicios a través de los web services, API (interfaz de programación de aplicaciones) o interfaz de usuario.

Una vez que se han creado o capturado los datos y establecido la forma en que se van a almacenar, **resulta clave que los datos puedan moverse, transmitirse y estar disponibles según las necesidades de la empresa, tanto entre sistemas internos como externos.**

Construir una red de transmisión de datos requiere del uso conjunto de **software** (sistema operativo de red y software de aplicación o de programas específicos), **hardware**, **dispositivos de usuario final y de red** (servidores, almacenamiento en red) y **protocolos de comunicaciones**.

"Resulta clave que los datos puedan moverse, transmitirse y estar disponibles según las necesidades de la empresa"

La gobernanza de los datos en esta fase implica también establecer **protocolos para poder garantizar la seguridad de la información corporativa** como, por ejemplo, el uso de PKI (infraestructura de clave pública), firmas digitales o gestión de derechos digitales.

Durante esta fase, se debe resolver a nivel tecnológico:

- Definir un esquema para almacenar y recuperar la información.
- Sistema que se utiliza para almacenar los datos.
- Dónde guardar la información.
- Definir los componentes básicos de las redes de transmisión de datos (software hardware y protocolos).
- Políticas de seguridad y protección de los datos (copias de seguridad, control de accesos).
- Políticas y sistemas para el archivo y almacenamiento de la información del negocio a largo plazo (Preservación).

3. Gestión, acceso y compartición

En esta fase del ciclo de vida de los datos, hay que dar respuesta a cómo conseguir la información adecuada para las personas o sistemas adecuados en el dispositivo correcto. Es decir, **integrar el sistema de gestión de los datos con el resto de aplicaciones y entornos organizativos.**

La transformación de los datos en información útil para el negocio abarca distintas disciplinas y subsistemas:

- **Gestión Documental:** software que controla y organiza los documentos de una organización. Incorpora la captura de documentos y contenidos, flujos de trabajo con la documentación (revisión, aceptación, publicación, etc.), repositorios de documentos, sistemas de salida y sistemas de recuperación de la información.

- **Gestión de archivos digitales:** gestión de ficheros electrónicos que usualmente contienen imágenes, vídeos, dosieres y otros materiales de valor para el negocio.

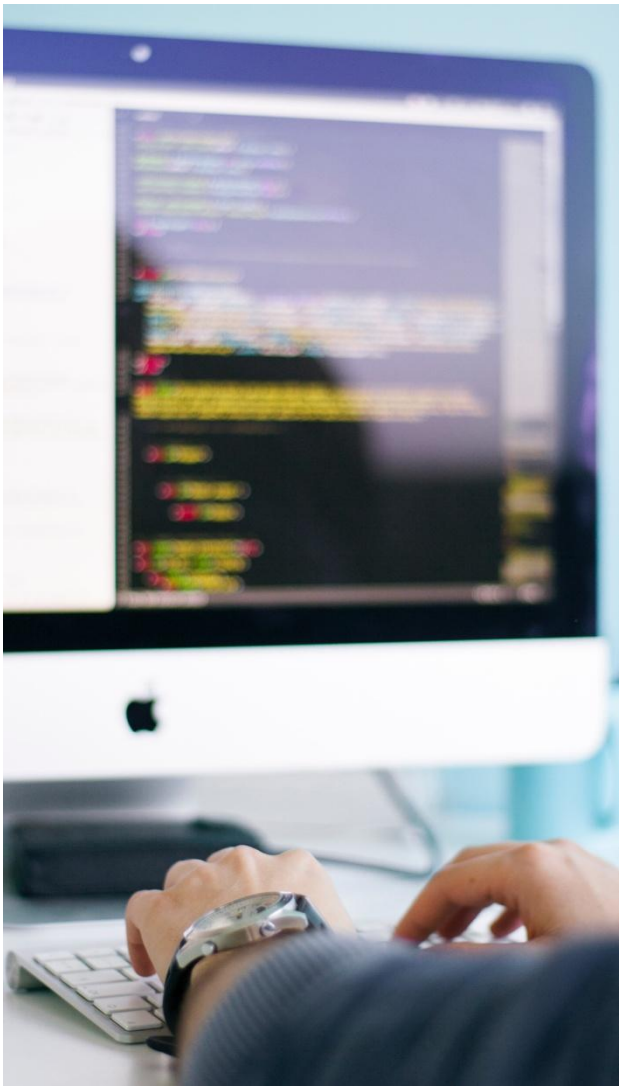
- **Gestión de contenidos web:** tecnología que se ocupa de la creación, revisión, aprobación y publicación del contenido basado en Web.

- **Colaboración:** herramientas que permiten que varios usuarios trabajen en el mismo contenido en un entorno común.

- **Flujos de trabajo o Workflows / BPM (Gestión de procesos de negocio):** automatización de procesos de negocio, en su totalidad o en parte, en los que los documentos, la información o las tareas se pasan de un participante (o sistema) a otro para la acción, de acuerdo con un conjunto de reglas. Un proceso de negocio es un conjunto lógicamente relacionado de flujos de trabajo, pasos de trabajo y tareas que proporcionan un producto o servicio a los clientes. BPM es una mezcla de gestión de procesos / flujo de trabajo con tecnología de integración de aplicaciones.

- **Gestión de registros**: permite a una empresa asignar un ciclo de vida específico a partes individuales de información corporativa, desde la creación, la recepción, el mantenimiento y el uso hasta la disposición final de los registros. Un registro no es necesariamente el mismo que un documento. Todos los documentos son registros potenciales, pero no viceversa. Un registro es esencial para el negocio. Los documentos son contenedores de "información de trabajo". Los registros son documentos con valor probatorio.

Para la compartición de los datos, se usan distintas tecnologías de transformación (COLD/ERM, XML, PDF, Compresión, Sindicación) que permiten la distribución de la información a través de Internet, Intranet, Extranet, E-mail, dispositivos móviles o incluso papel.



4. Análisis, información e inteligencia de negocio

La cuarta fase del Ciclo de Gestión de los datos, implica el **análisis, explotación y uso de los datos aplicado a la toma de decisiones de negocio.**

Business Intelligence o Inteligencia de negocio se refiere a **la creación de conocimiento a partir del análisis de los datos de una empresa** mediante una estrategia en la que se combinan datos, aplicaciones, productos/servicios, tecnologías y una determinada arquitectura a nivel técnico. De forma que, el análisis y comprensión de los datos actuales facilite la **previsión de eventos futuros y apoye una toma de decisiones basada en evidencias.**

A su vez, se debe garantizar **el movimiento y la circulación de los datos** entre los diferentes sistemas o aplicaciones informáticas, para **que la información sea accesible**, así como que la usabilidad del sistema para su uso no requiera de elevados conocimientos técnicos por parte de los usuarios finales.

La **Computación cognitiva** o sistemas ‘inteligentes’ además de procesar la información contenida en los datos, son capaces de interpretar la información para aprender, razonar, resolver problemas o realizar predicciones que ayuden a las organizaciones a tomar mejores decisiones a través del análisis de su información.

A través del análisis, información e inteligencia de negocio se pueden generar:

- Informes, consultas y alertas.
- Optimizaciones de procesos.
- Análisis estadísticos.
- Indicadores de gestión.
- Patrones o tendencias de compra o uso, ciclos de producción.
- Previsiones.
- Control de procesos y costes.

Caso práctico de aplicación de la Gestión del Ciclo de Vida de los Datos

Como hemos visto en los anteriores epígrafes, existe una **amplia variedad de soluciones tecnológicas que permiten gestionar cada fase del ciclo de los datos**. Si bien, la combinación de tecnología, datos y procedimientos para convertir la información en un activo del negocio requiere de un análisis exhaustivo y el conocimiento profesional de las tecnologías y técnicas adecuadas.

En Clase10 somos especialistas en soluciones de Business Intelligence. A continuación. explicamos el caso práctico de la compañía Tenneco Automotive, S.A. donde una aplicación adecuada de tecnología y procedimientos permitió a la organización importantes ahorros en tiempo y costes y optimizar sus procesos.

Tenneco Automotive Ibérica, S.A.

La empresa Tenneco Automotive Ibérica, S.A. posee diferentes centros de trabajo en los que se fabrican diversas piezas para ser posteriormente ensambladas, dando lugar a un producto final.

Esta empresa utiliza una serie de vehículos motorizados con vagones denominados ‘recogepedidos’, los cuales se encargan de recoger y trasladar dichas piezas para su ensamblaje. El reto al que se enfrentó Tenneco era el de optimizar la producción mediante el estudio de las rutas, turnos, gestión de la plantilla, tiempo, etc.

Solución aplicada

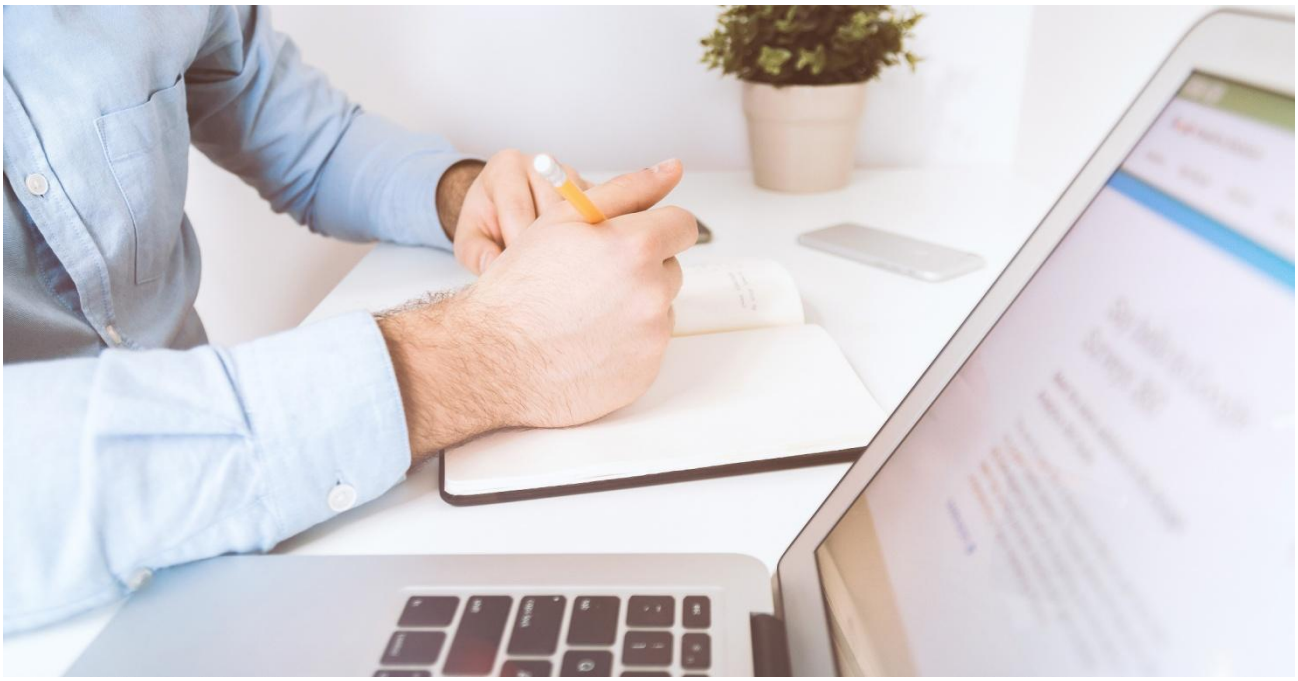
Tras analizar la situación, C10 adoptó la siguiente solución tecnológica: la incorporación para cada pieza fabricada de **etiquetas electrónicas** y la instalación de lectores de **tecnología RFID** para los vehículos.

Aquí es donde empieza la Gestión del Ciclo de Vida de los Datos en este caso de éxito. Los recogepedidos van circulando y realizando sus tareas como de costumbre. Gracias a los lectores y a las etiquetas electrónicas se crean una serie de datos para la empresa Tenneco.

Tras pasar por las diferentes fases vistas anteriormente relacionadas con los datos y su transmisión, almacenamiento, seguirdad, gestión, etc., se llega al momento en el que dichos datos son explotados con la idea de poder **obtener un conocimiento aplicable que mejore el rendimiento productivo.**

Aplicando la tecnología RFID y el análisis de la información, Tenneco obtuvo un **ahorro anual estimado de unos 100.000 €, una reducción del tiempo de manejo de las distintas piezas del 20% y un incremento del 30% en los ciclos por turno.**

"Tenneco incorporó etiquetas electrónicas para cada pieza fabricada e instaló lectores de tecnología RFID para los vehículos"



Contacta con uno de nuestros consultores de Clase10 y estudiaremos tu caso específico para aplicar en tu organización una solución que te permita convertir la información en conocimiento.

Quiero hablar con un consultor de C10

C10
EDIFICIO CLASE10
Parque Tecnológico
C/ Leonardo da Vinci, 30
46980 Paterna
Valencia (España)

Tel: 961 366 069
Fax: 961 366 219
Email: info@clase10.com

C10

www.clase10.com