

UF1886: Administración del gestor de datos
en sistemas ERP-CRM

Elaborado por: Juan José Gutierrez Cañizares

Edición: 5.0

EDITORIAL ELEARNING S.L.

ISBN: 978-84-16557-53-0

No está permitida la reproducción total o parcial de esta obra bajo cualquiera de sus formas gráficas o audiovisuales sin la autorización previa y por escrito de los titulares del depósito legal.

Impreso en España - Printed in Spain

Presentación

Identificación de la Unidad Formativa

Bienvenido a la Unidad Formativa **UF1886: Administración del gestor de datos en sistemas ERP-CRM**. Esta Unidad formativa pertenece al Módulo Formativo **MF1214_3: Administración de sistemas de planificación de recursos empresariales y de gestión de relaciones con clientes**, que forma parte del Certificado de Profesionalidad **IFCT0610: Administración y programación en sistemas de planificación de recursos empresariales y de gestión de relaciones con clientes**, de la familia de **Informática y Comunicaciones**.

Presentación de los contenidos

La finalidad de esta Unidad Formativa es enseñar al alumno a administrar el almacén de datos para mantener y asegurar su funcionamiento, de acuerdo a los criterios establecidos por la organización y a las recomendaciones del fabricante del producto.

Para ello, en primer lugar se analizará la gestión de mantenimiento en almacén de datos, la evaluación del rendimiento y la resolución de problemas en el almacén de datos. También se estudiará el transporte de componentes entre entornos de desarrollo, prueba y explotación en almacén de datos, los errores en la ejecución del transporte y los procesos de extracción de datos en sistemas de almacén de datos. Por último, se profundizará en las incidencias en el proceso de extracción de datos.

Objetivos de la Unidad Formativa

Al finalizar esta Unidad Formativa aprenderás a:

- Describir el funcionamiento y gestionar el sistema de transporte de objetos y componentes entre los diferentes entornos de desarrollo, prueba y explotación, para mantener el flujo de modificaciones y su verificación antes de ser liberadas para su utilización por el usuario final, según requerimientos técnicos y funcionales.
- Describir el funcionamiento y lanzar periódicamente los procedimientos de extracción de datos que alimentan de datos los sistemas de ERP, CRM y almacén de datos para mantener la actualidad de su información, siguiendo especificaciones técnicas.

Índice

UD1. Gestión de mantenimiento en almacén de datos	9
1.1. Procesos de los Sistemas de Gestión de Datos.....	11
1.1.1. Definición de la estructura de datos	16
1.1.2. Definición de la estructura de Procesos.....	42
1.1.3. Integración de Procesos.....	64
1.2. Parámetros de los Sistemas que Influyen en el Rendimiento ..	66
1.2.1. Definición de Parámetros de Configuración	67
1.2.2. Optimización de recursos de software y tiempos de ejecución.....	91
UD2. Evaluación del Rendimiento	103
2.1. Herramientas de monitorización y evaluación del rendimiento: características y funcionalidades	105
2.1.1. Definición de las funcionalidades que se desea evaluar	189
2.1.2. Definición de las herramientas para evaluar el rendimiento del sistema ERP.....	191

UD3. Resolución de problemas en el almacén de datos	205
3.1. Trazas del Sistema (logs)	207
3.1.1. Definición de sistemas para trazar los procesos entre los sistemas ERP, CRM y Almacén de Datos.....	236
3.2. Incidencias: identificación y resolución	251
3.2.1. Definición de los procesos para identificar incidencias, control y resolución	265
3.2.2. Documentos de las tareas e incidencias realizadas ...	277
UD4. Transporte de componentes entre entornos de desarrollo, prueba y explotación en almacén de datos	287
4.1. El sistema de transmisión de información.....	289
4.1.1. Características en la transmisión de datos.....	291
4.1.2. Herramientas que intervienen en la transmisión de datos, configuración de parámetros.....	315
4.2. Entornos de Desarrollo.....	351
4.2.1. Definición de los entornos de desarrollo, pruebas y explotación	353
4.2.2. Gestión del transporte entre los entornos de desarrollo, pruebas y explotación.....	365
UD5. Errores en la ejecución del transporte.....	373
5.1. Tipos de errores.....	375
5.1.1. Verificación de la sintaxis	378
5.1.2. Identificación de las incompatibilidades entre componentes	387
5.1.3. Documentación de errores encontrados	397
5.2. Resolución de Problemas.....	405
5.2.1. Utilización de manuales para la resolución de problemas .	408
5.2.2. Documentación de las soluciones aportadas	416

UD6. Procesos de extracción de datos en sistemas de almacén de datos	425
6.1. Características y funcionalidades	427
6.1.1. Definición de la estructura que interviene en los proce- sos de extracción de datos	448
6.1.2. Tratamiento de la información y transformación de es- tos datos para facilitar la toma de decisiones.....	453
6.1.3. Interpretación de Resultados	456
6.2. Procedimientos de ejecución.....	459
6.2.1. Definición de procedimientos de extracción de datos entre los sistemas ERP, CRM y la base de datos.....	466
UD7. Incidencias en el proceso de extracción de datos.....	545
7.1. Trazas de ejecución	547
7.1.1. Establecimiento de trazas para estudiar incidencias...	554
7.2. Resolución de incidencias.....	581
7.2.1. Solución y documentación de incidencias.....	582
Supuesto práctico 1	593
Supuesto práctico 2	605
Glosario	611
Soluciones	619

Área: informática y comunicaciones

UD1

Gestión de
mantenimiento en
almacén de datos

- 1.1. Procesos de los Sistemas de Gestión de Datos
 - 1.1.1. Definición de la estructura de datos
 - 1.1.2. Definición de la estructura de Procesos
 - 1.1.3. Integración de Procesos
- 1.2. Parámetros de los Sistemas que Influyen en el Rendimiento
 - 1.2.1. Definición de Parámetros de Configuración
 - 1.2.2. Optimización de recursos de software y tiempos de ejecución

1.1. Procesos de los Sistemas de Gestión de Datos

Existen muchas definiciones de almacén de datos. Se puede considerar un almacén de datos como una sala donde se guardan expedientes en formato de papel. Aunque en el contexto del procesamiento de datos mediante almacén de datos o data warehouse tiene un significado un poco más complejo que el de un simple almacén.

En el ámbito del procesamiento de datos por equipos informáticos un Data Warehouse implica una colección de datos orientada a un determinado ámbito, es decir a una organización, empresa o sector público. Cuando esta definición indica que está orientada “a un determinado ámbito”, esto quiere decir que implica ciertas características:

- Integrado en el sistema de gestión de la organización.
- No Volátil.
- Variable en el tiempo.
- Consultable.
- Organizado por temática.

De tal modo que cumpliendo todas estas características se consigue un sistema que no solo almacena sino que ayuda a la toma de decisiones de la organización o entidad que lo utiliza.

Resumiendo, se trata de un informe o expediente completo de una organización, algo mucho más extenso y completo que la simple información referente a transacciones o a operatividad. De este modo se consigue un sistema de base de datos que favorece, no solo el análisis, sino la divulgación eficiente de los mismos.

Otra de las particularidades de este sistema es que no debe usarse con datos actuales o que estén implicados en operaciones o transacciones en curso.

Los Data Warehouses dividen su información en unidades lógicas más pequeñas dependiendo de la organización de la empresa o entidad. Esto se debe a la gran cantidad de datos que gestionan.

Existe una definición de uno de los más conocidos autores de esta temática, Ralph Kimball, que define el almacén de datos cómo:

"Una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis"[...] "la unión de todos los data marts de una entidad".

Los data marts se consideran pequeños data warehouse que están centrados en un tema, área de negocio o ámbito de aplicación específico dentro de una organización. En síntesis los Data marts son subconjuntos de datos de un data warehouse para áreas específicas.

Estos data marts tienen unas características, entre las cuales destacan:

- Usuarios limitados.
- Área específica.
- Tiene un propósito específico.
- Tiene una función de apoyo.

Muchas referencias a un almacén de datos incluyen una definición más amplia en la que se agregan herramientas para:

- Extraer.
- Transformar.
- Cargar datos.
- Herramientas para un análisis de inteligencia empresarial.
- Herramientas para gestionar y recuperar los metadatos.

Al integrar en una organización un almacén de datos lo que se pretende es contener datos de diversa utilidad. En resumen se puede indicar que se utiliza como una consigna de datos para su posterior uso y transformación en información útil para el usuario.

Un almacén de datos debe cumplir una función vital, entregar la información correcta a la gente indicada en el momento y formato óptimo.

El almacén de datos da un gran soporte a sus usuarios que deseen usar uno de los siguientes sistemas:

- Sistemas de Soporte a Decisiones (DSS).
- Sistemas de información ejecutiva (EIS).
- Herramientas para consultas o informes.

Gracias al uso de estas herramientas los usuarios finales pueden realizar fácilmente consultas sobre sus almacenes de datos evitando la interacción directa con los datos, no comprometiendo así ni la integridad ni el formato de estos.

Basándonos en la funcionalidad del Data Warehouse, se ha de tener en cuenta la vital importancia de dos de sus funciones:

- **Integración de los datos.** Los datos no suelen provenir de una sola base de datos, sino que son de varias bases de datos distribuidas por las diferentes unidades de la organización. Además se ha de tener en cuenta, que con frecuencia la estructura de estas bases de datos son parcial o incluso totalmente diferentes. Es decir que con frecuencia las fuentes son heterogéneas. En estos casos se hace extremadamente importante el indicar tanto una descripción global como un análisis comprensivo de toda la organización final en el almacén de datos.
- **Separación de los datos.** Teniendo en cuenta la importancia de la separación de los datos de operación y transacción diaria respecto de los datos almacenados para otros propósitos como:
 - Divulgación.
 - Ayuda en la toma de decisiones.
 - Análisis y operaciones de control.

Estos datos, diarios y de otros propósitos, no han de coincidir sobre la misma base de datos.

En el funcionamiento de estos almacenes de datos incluye un volcado periódico de datos desde los distintos sistemas de planteamiento de recursos (ERP) y de otros sistemas en los que se almacenen o gestionen.

Antes del volcado de datos desde los diferentes entornos de la empresa u organización, estos deben tratarse para homogeneizarse con los datos del almacén

de datos donde se va a integrar. Este se suele realizar mediante herramientas de extracción, transformación y carga, denominadas normalmente ETL.

Estas herramientas ETL revisan los datos primarios realizando un proceso de transformación previo al volcado al almacén de datos.

Elementos y Herramientas que Forman el Almacén de Datos

Un almacén de datos intervienen una amplia variedad de elementos, incluyendo un más que importante factor humano. No obstante, para la comprensión de la funcionalidad de estos sistema, se han resumido en tres factores fundamentales: Metadatos, Funciones ETL y Middleware.

Estos elementos principales que componen un almacén de datos se pueden resumir en:

- **Metadatos.** Su principal definición es la de “datos acerca de datos”. Esto quiere decir que son datos que describen la estructura de otros datos que se van a almacenar y de sus relaciones.

Estos indican las tablas de las que se compone una base de datos, que columnas o registros tiene cada tabla y que tipo de datos pueden almacenar.

En el ámbito de los data warehouse, el metadato juega un papel fundamental dado que recoge todas las definiciones de la organización y el almacén de datos. Contiene toda la información de:

- Tablas.
- Entidades y Relaciones.
- Columnas de tablas.
- Relaciones entre tablas.
- Jerarquías y Dimensiones de datos.

- **Funciones ETL.** Los procesos ETL son de vital importancia e implican las siguientes operaciones:

- **Extracción.** Operación de recopilación de la información deseada a partir de los datos almacenados en fuentes diferentes al almacén de datos.
- **Transformación.** Operación sobre los datos para que puedan ser cargados en el Almacén de datos o se puedan migrar de éste a otra base de datos.

- **Carga.** Operación de Almacenamiento de los datos en la base de datos en el almacén de datos.
- **Middleware.** Se refiere al software de conectividad entre aplicaciones distribuidas sobre plataformas heterogéneas. Se puede incluir en las definiciones de capas API.

Su función en el contexto de los almacenes de datos es asegurar la conectividad entre los componentes de software, incluyendo de aplicación y sistema, que están implicados en el sistema de gestión del almacén de datos.



Middleware se denomina también “logical de intercambio de información entre aplicaciones” o “interlogical”. Este software funciona como “Capa de abstracción” y se sitúa entre “Capa de Aplicación” y las capas inferiores.

En el diseño de un Data Warehouse se requieren herramientas para poder migrar y transformar los datos a este almacén y una vez en funcionamiento se requiere el manejo de grandes volúmenes de información.

Se ha de tener en cuenta que la arquitectura de cada Sistema de Almacén de Datos dependerá en gran medida de los datos que se pretendan almacenar y la estructura que se les quiera dar. Debido a esto, los datos deben ser repartidos entre diversos data marts.

Es por estos factores que para el diseño de un data warehouse es necesario tener el amplios conocimientos de algunos temas generales de la organización o empresa, los cuales se describen a continuación:

- **Situación inicial.** Es evidente que las soluciones propuestas para el almacenamiento en Data Warehouse, como cualquier solución de software para la empresa, deben estar orientadas a las necesidades del negocio.

Aunque en este punto se ha de incluir que deben estar en concordancia con la arquitectura y programas actualmente existentes en el organigrama de la empresa.

- **Tipo de negocio.** En este punto se ha de tener en cuenta no solo el tipo de negocio que existe en este momento. Sino que se ha de tener un conocimiento exacto de los procesos pasados, presentes y estrategias de futuro, con el fin de poder homogeneizar toda la actividad.
- **Entorno operativo.** Consideración del Hardware, Aplicaciones y Herramientas que se usan en el día a día de la empresa.
- **Usuarios.** Los usuarios serán los que accedan al sistema, lo implementen y, al fin y al cabo lo usen a diario. Con lo cual se debe tener en cuenta lo que los usuarios necesitan para el control del Almacenamiento de Datos.

Es evidente que al tener en cuenta tantos factores, estos sistemas tienen vinculados algunas ventajas e inconvenientes.

Ventajas
<ul style="list-style-type: none"> • Hacen más sencillo el acceso a una gran variedad de datos a los usuarios finales. • Ayudan al funcionamiento de las aplicaciones de los sistemas de apoyo, sobre todo en generación de informes.
Inconvenientes
<ul style="list-style-type: none"> • Pueden suponer costes de mantenimiento elevados. • Tienen una rápida obsolescencia en el actual desarrollo tecnológico.

1.1.1. Definición de la estructura de datos

Cuando se define una estructura de datos también implica que existe datos estructurados para formarla. Es decir, que está formada por una serie de datos simples o elementales y relacionados entre ellos.

En el entorno de los Sistemas de Gestión de Datos, al referirse a Estructuras de Datos se hace referencia a un conjunto de datos que poseen el mismo nombre. Estos datos tienen también la característica de que pueden ser ordenados por su organización y por las operaciones que se definen en el.

Las estructuras de datos se pueden clasificar en:

- **Homogéneas.** Todos los datos simples o elementales que las forman son del mismo tipo.
- **Heterogéneas.** Todos los datos simples o elementales que las forman son de varios tipos o formatos.

Los datos que formaran la estructura en un programa deben estar determinados de antemano. Esto es debido a que para su posterior uso se debe tratar y almacenar por un “traductor”. En caso de que los datos sean elementales son estos los que determinan el espacio que se utiliza en la memoria, en caso de datos estructurados puede o no ser así.

Datos Simples	
Estándar	<ul style="list-style-type: none"> • Entero • Real • Caracteres • Lógico

Estos datos simples, se incorporarán a una estructura para formar la estructura de datos.

Estructura de Datos		
Estáticas	<ul style="list-style-type: none"> • Vectores y Matrices • Registros • Archivos 	
Dinámicas	<ul style="list-style-type: none"> • Lineales 	<ul style="list-style-type: none"> • Pilas • Colas • Listas Enlazadas
	<ul style="list-style-type: none"> • No Lineales 	<ul style="list-style-type: none"> • Árboles • Grafos

Como se estudiará en detalle más adelante, una estructura de datos puede ser:

- **Estática.** Esta definición corresponde a una estructura de datos que siempre ocupará el mismo espacio en la memoria. En este tipo de organización el volumen que ocupará la estructura se define antes de que el programa se ejecute. Estas estructuras arrastran un inconveniente, el tamaño reservado no puede ser modificado durante la ejecución.
- **Dinámica.** Esta definición corresponde a una estructura de datos que realiza una asignación dinámica de memoria. Este tipo de estructuras no existen limitaciones en cuanto al tamaño de memoria ocupada. No obstante, para utilizar este tipo de estructuras necesitamos usar un tipo de dato específico llamado puntero.

Estructura de Datos Estáticas

Estas estructuras de datos estáticas que ocupan siempre el mismo espacio en la memoria se subdividen en tres tipos:

- Puntero.
- Cadena.
- Array.

Un Puntero es un dato que indica la posición de otro. Es decir, es un dato cuya utilidad principal es indicar otro dato. Tienen una especial utilidad en la estructura de datos ya que son el nexo de unión entre datos. Existen dos punteros especialmente importantes, el de que indica el principio y el final de la estructura.



Si por funcionalidad se requiere un puntero que no señale a ningún dato se dice que este puntero tiene un valor nulo, o puntero nulo.

Una Cadena se puede definir como una secuencia de caracteres que se interpretan como un único dato. Es decir se corresponde con varios caracteres que pueden incluir espacios y tienen un longitud fija o longitud variable.

Esta longitud de la cadena se indica mediante dos factores:

- Número de Caracteres que Contiene.
- Carácter Fin de cadena.

Las operaciones que se puede realizar sobre datos de tipo cadena son cuatro:

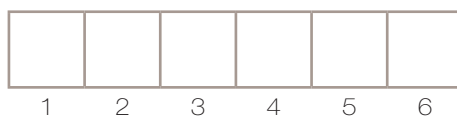
- **Concatenación.** Operación consistente en formar una cadena a partir de otras dos mediante la unión de los caracteres de ambas.
- **Extracción de Subcadena.** Operación consistente en formar una cadena a partir de un tramo de otra ya existente.

- **Comparación de cadenas.** Operación consistente en comparar dos cadenas. Se considera menor aquella en que el primer carácter en que difieren ambas sea menor.
- **Obtención de la longitud.** Operación consistente en la obtención de la longitud de una cadena, da como resultado un dato de tipo entero cuyo valor es el número de caracteres que contiene la cadena.

Un Array también se denomina formación o matriz e incluso arreglo, es una de las estructuras de datos más usual. Existen muchos de los lenguajes de programación que usan esta denominación, prácticamente todos.

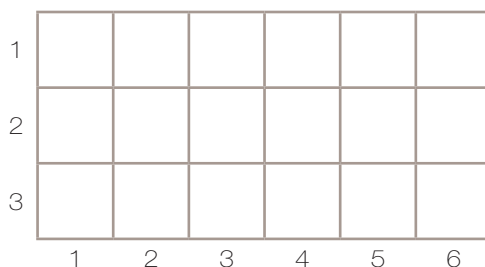
Un Array es una estructura de datos que está formada por una cantidad de datos finita del mismo tipo. Cada uno de los datos tiene asociado uno o varios índices que indica la posición exacta de un dato en un Array.

Se puede hacer el símil de un Array con una estructura de celdas donde se pueden almacenar valores. Por ejemplo suponiendo una de un solo índice que tomara valores del 1 al 6, quedaría como en el esquema:



Estructura de Datos Estáticas

Ampliando el ejemplo de Array anterior se puede observar en el siguiente esquema. En esta matriz cada elemento está designado por un par ordenado de números que representaran la filas y columnas. En la estructura de procesos detallará la operatividad de este tipo de Arrays.



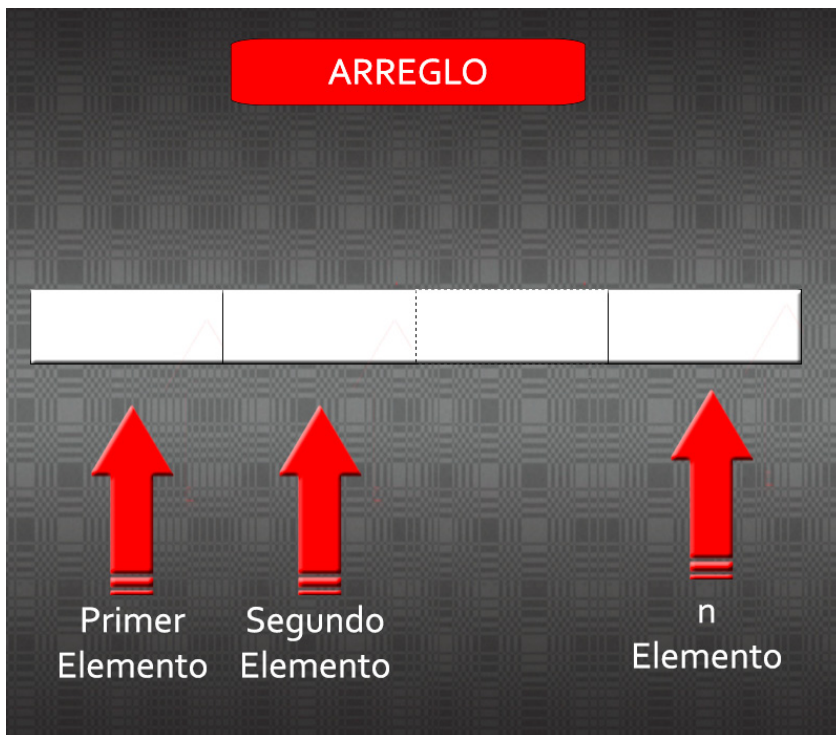
Los índices del Array son los que denominan las dimensiones del Array.

Estructura de Datos Estáticas. Arrays

Se puede definir “Arreglo” es un conjunto de datos o una estructura de datos homogéneos que se encuentran ubicados en forma consecutiva en la memoria RAM, en resumen, sirve para almacenar datos en forma temporal.

Los Arrays se pueden definir por como una colección finita, homogénea y ordenada de elementos.

- **Finita.** Todo array, al ser un conjunto de datos homogéneos, tiene un límite. En esencia quiere decir que puede determinarse cuál será el número máximo de elementos que podrán formar parte del array.
- **Homogénea.** Todos los elementos de un array son del mismo tipo o naturaleza pero nunca una combinación de distintos tipos. Es decir, todos los elementos son enteros o caracteres o lógicos
- **Ordenada.** Se debe determinar de manera clara cuál es el orden de los elementos, desde el primero hasta el enésimo.



Una de las características más relevantes de los arrays es poder determinar el orden de cada elemento