

# HERRAMIENTAS INFORMÁTICAS PARA GESTIONAR EL CONOCIMIENTO EN LAS EMPRESAS.

Autor : Ixone Alonso Sanz

[eupalsai@bs.ehu.es](mailto:eupalsai@bs.ehu.es)

Laboratorio de Informática Aplicada  
Facultad de Ciencias Económicas y Empresariales  
*Universidad del País Vasco (U.P.V/E.H.U)*

Palabras clave: Datos, Información, Conocimiento, Datawarehouse, Datamining, Olap, CRM

## ABSTRACT

Nos encontramos ante una nueva era económica denominada nueva economía, era de la información, sociedad del conocimiento, caracterizada porque sus fuentes principales de riqueza son la *información* y el *conocimiento*. Antes de establecer la diferencia entre ambas, se hace preciso mencionar que la materia prima de la información son los *datos*. Estos constituyen el nivel más bajo de los hechos conocidos y no tienen un significado intrínseco por sí mismo, pero cuando son ordenados, agrupados, analizados e interpretados se convierten en *información*.

Esta a su vez constituye la materia prima del *conocimiento* ya que éste es la combinación de *información*, contexto y experiencia.

Las empresas han de disponer de las herramientas adecuadas para poder almacenar los *datos*, transformar éstos en *información* y así mismo ésta en *conocimiento*.

El *Datawarehouse* es la plataforma que concentra toda la información que es de interés para la organización, constituyendo el núcleo sobre el que se van a aplicar las *herramientas informáticas de ayuda a la decisión* como el *Datamining*, *Olap*, *CRM*, que van a permitir que el usuario final pueda tomar decisiones y convertir la información en conocimiento.

## 1. LA NUEVA ECONOMÍA

La idea de la *nueva economía* ha venido de la mano del sector electrónico y de las telecomunicaciones, cuya intromisión económica y social es determinante para el crecimiento, estructuras y formas de la empresa.

La globalización de los mercados, la diseminación de las tecnologías de la información y el desmantelamiento de las jerarquías nacidas a mediados del siglo pasado, son los rasgos esenciales de una nueva era económica caracterizada porque sus fuentes principales de riqueza son la **información y el conocimiento** frente a los tradicionales recursos naturales y trabajo físico. No obstante, la *tierra, el capital y el trabajo*, no han desaparecido, ni lo harán en el sentido de su existencia, pero los cambios tecnológicos que ha propiciado esta nueva coyuntura implican una reconceptualización de estos tradicionales recursos, al conformarse nuevas estructuras organizacionales en función del nuevo **conocimiento** que la **información** produce.

Esta *sociedad de la información* también se caracteriza por la riqueza de medios que pone a nuestro alcance para localizar **información**, almacenarla, procesarla y transportarla. Son las nuevas tecnologías y en especial *Internet* las que están posibilitando que tengamos un acceso a tanta información.

Lo que parece claro es que el desarrollo revolucionario de las nuevas tecnologías de la *sociedad de la información* ha provocado una profunda transformación en las **empresas** de todo tipo.

La empresa es ahora un **centro de innovación**, entendiendo por innovación el proceso de transformación del conocimiento en riqueza, siendo la palanca que impulsa y hace posible el desarrollo económico sostenible. La empresa de la *revolución de la información* innova en base a un proceso de acumulación de capitales intangibles, tecnología y recursos humanos cualificados.

Esta cualificación de los recursos humanos pasa por la *creatividad y el ingenio*, dos características importantes de los recursos humanos, ya que pueden ser fuentes de ventaja competitiva.

*Internet* constituye un factor importante para las empresas y no en vano es el motor de la *nueva economía* pero esto no implica que el único objetivo de las empresas sea vender por Internet, esto es, no basta con hacerse con un *punto.com*.

La *nueva economía* va más allá que todo eso, la palabra clave es la **propuesta de valor**, que se debe hacer a los clientes. Esto deja entrever que las reglas del negocio cambian y a partir de ahora *todo debe girar alrededor del cliente*.

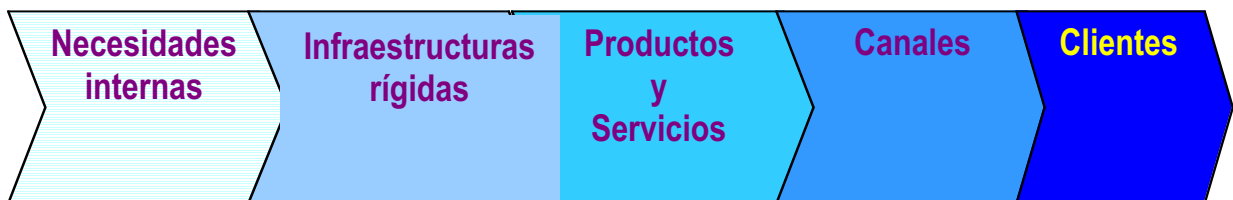
Pero para ello es necesario que la empresa disponga de la **información** suficiente para poder satisfacer las necesidades de los clientes. Por tanto, el *producto información* aparece como una solución idónea a cada problema o necesidad.

La empresa lanzará **productos y servicios que sean novedosos** pero teniendo en cuenta lo que demandan sus clientes. Por eso en las empresas *e-bussines* estos productos que ofrecen las empresas se suelen denominar "**produccio**" porque son productos y servicios *recubiertos de una capa informacional*.

En la vieja economía las empresas lanzaban sus productos a partir de la información que respondía a las propias necesidades de la empresa. En este nuevo escenario, las empresas lanzan sus productos estableciendo previamente un contacto informacional con el cliente para saber si se ajusta a sus necesidades. Por tanto, ya no basta con tener los mejores productos o servicios al mejor precio, ni los canales de distribución más eficientes, las reglas de juego han cambiado y son los *clientes quienes dictan las normas*.

Esto supone un cambio de pensar en las empresas, ya que en los negocios tradicionales se pensaba *primero en la propia empresa* y lo único que interesaba era mantener la cifra de ventas, con independencia de la opinión de los clientes. En la nueva economía se *piensa primero en el cliente* antes que en la empresa. Esto se ve reflejado también en la **cadena de valor** de las empresas, produciéndose un giro total de la misma.

En la *vieja economía* se producía el siguiente proceso hasta llegar al cliente final



En la *nueva economía*, la cadena de valor se invierte



La entrada de la *nueva economía* se traduce también en modificaciones profundas en el seno de la **estructura de costes** de las empresas. El ejemplo más significativo es el de los *costes de distribución*: en el caso de un billete de avión, dichos costes representan una media de 8\$ cuando se compran en una ventanilla, mientras que bajan a 1\$ cuando ese billete se compra a través de Internet. Así, el e-business tendrá un impacto sobre los *costes de inventario* y los *stocks*.

En las empresas de la *nueva economía* los **trabajadores del conocimiento** se convierten en el *principal activo* ya que cada uno de ellos va a aportar valor con sus conocimientos a la empresa, es decir, va a ser capaz de desarrollar y mejorar sus competencias en el marco de la empresa.

Así, el mayor reto de las empresas es el integrar en sus sistemas operativos a los *trabajadores del conocimiento*, indispensables para la permanencia del proceso innovador y para la supervivencia de la empresa. El trabajador de la *era de la información* tiene una función nodular en la red del conocimiento, su trabajo es en sí mismo un producto que enriquece la empresa, pero que al mismo tiempo contribuye a la realización personal del trabajador.

Esto es, las empresas son lugares en los que **crear nuevo conocimiento no es una actividad especial**, sino que es una forma de comportarse, incluso de ser, y en donde todas las personas son *trabajadores del conocimiento*. Así se puede observar que el *poder* está en el *cliente* y en el *profesional*. En el *cliente*, porque la atención de la empresa debe centrarse en primer lugar en él y en el *recurso humano*, ya que debe ser capaz de crear **conocimiento** a partir de las grandes masas de **información** que posee y así generar valor a la organización.

## **2. DATOS, INFORMACIÓN Y CONOCIMIENTO.**

Una vez examinados algunos de los efectos que sobre las empresas tiene la *nueva economía*, es preciso establecer las diferencias existentes entre los recursos más valiosos que disponen las empresas de la actualidad: **información y conocimiento**.

Pero antes de hacer referencia a la **información** se hace preciso mencionar a los **datos**, ya que éstos son el nivel más bajo de los hechos conocidos y sin ellos no sería posible configurar la información al constituir éstos la materia prima de la misma. Así, los **datos**

constituyen el nivel más bajo de los hechos conocidos y por sí mismos no dicen nada sobre el porqué de las cosas.

Describen únicamente una parte de lo que pasa en la realidad, no proporcionan juicios de valor o interpretaciones y, por tanto, no son orientativos para la *acción*, sino que van asociados a un *objeto*.

Tienen dos características distintivas. Por un lado, los datos quedan perfectamente definidos sin posibilidad de confusión por un conjunto de símbolos (letras y números) y por otro, son contrastables, es decir, es posible determinar si un dato es cierto o no.

Los **datos** no tienen un significado por sí solos pero en el momento que son *ordenados, agrupados, analizados e interpretados*, se convierten en **información**.

Luego ésta son datos *dotados de significado* ya que se encuentran envueltos en un determinado *contexto*.

El grado de participación humana es importante ya que es el *recurso humano* el que se va a encargar precisamente de *dotar de sentido y significado* a los **datos** capturados en las máquinas y convertirlos en **información**.

Mediante un ejemplo muy sencillo se puede reflejar la diferencia entre **datos** e **información**. Si hay *100 unidades* de un determinado producto en el almacén de una empresa, esto en sí carece de relevancia. Ahora bien, si por término medio suele haber 1000 unidades, es obvio que *quedan pocas unidades en el almacén y hay que comprar más*. Esto es, el *dato inicial* al ser contextualizado es cuando se ha podido extraer una conclusión y por tanto transformarlo en *información*.

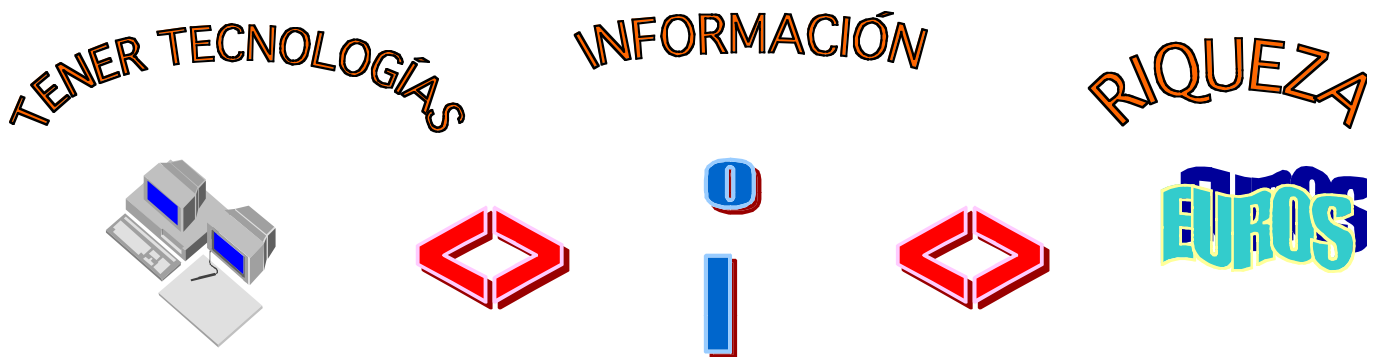
No es suficiente con ser capaces de dotar de significado a los **datos** para obtener **información**, porque de esta manera se acumularán grandes cantidades de información que las empresas no serán capaces de asimilar. El reto consiste en que los individuos de la empresa sean capaces de seleccionar aquella información que es *relevante* para la misma según el entorno en el que se desenvuelve. Pero es lógico, ya que si la información no es comprensible ni relevante para el receptor o receptores pierde su valor informacional.

Normalmente ocurre que una información es de interés para el emisor y no para el receptor. En este caso se está produciendo una *transferencia* de información.

Pero es preciso que exista un proceso de *transacción* para que la información emitida sintonice con las necesidades de quien la va a recibir, de forma que las expectativas de este último se cumplan y el proceso informativo sea eficiente.

Las tecnologías van a facilitar la capacidad de *transferencia* de la información pero no contribuyen a optimizar la *transacción* de información porque éste sigue siendo un

proceso básicamente humano. Por tanto, el tener tecnologías de la información, no implica necesariamente tener información porque ésta tiene que ser captada y utilizada adecuadamente para generar valor.



Cuando la **información** es utilizada y puesta en el contexto o marco de referencia de una persona se transforma en **conocimiento**.

Resulta interesante aportar una definición de este activo intangible ya que ocupa el lugar central en las empresas informacionales.

Para Davenport Y Prusak el **conocimiento** es una mezcla de experiencia, valores, información y "saber hacer" que sirve como marco de referencia para la incorporación de nuevas experiencias e información y es útil para la acción.

Según la definición se puede establecer que el **conocimiento** es la mezcla de **información**, *contexto* y *experiencia*. Así el conocimiento difícilmente podría generarse sin información, lo que se deduce que la *materia prima* fundamental del **conocimiento** la constituye la **información**.

El recurso humano se hace indispensable para la creación de conocimiento ya que éste es la información procesada en la mente humana.

El **conocimiento** y la **información** van a ser tratados como recursos diferentes (de hecho lo son), pero *complementarios* ya que el conocimiento va mucho más allá en el sentido de que es más complejo que la información, está estructurado y tiene más dimensiones que ésta. Éstas se plasman básicamente en lo que va a recubrir al recurso información, esto es, el *contexto* y la *experiencia* aportadas por el recurso humano.

*Esto lleva a establecer algunas diferencias entre ambos recursos.*

La **información** es **inerte y estática**, producido probablemente porque no existe tanta manipulación humana en este recurso como en el **conocimiento**.

Éste es **dinámico** ya que *reside en las personas* y son precisamente éstas las que van generando una "tormenta" de ideas de gran utilidad que serán transmitidas entre las mismas.

Precisamente esta mayor comunicación entre las personas generado por la compartición de ideas, diversas interpretaciones y extracción de conclusiones, lleva a enmarcar al **conocimiento** como un recurso más **subjetivo** que la información.

De hecho, la **información** es un recurso **objetivo** ya que son simplemente datos (hechos objetivos) que se van a situar en un determinado contexto y el ser contextualizados no implica una pérdida de objetividad. Sin embargo, el conocimiento es un flujo de informaciones que cada ser humano *interpretará, valorará y aportará las ideas que crea oportunas*.

Esta *relación tan estrecha existente entre el conocimiento y las personas* es lo que hace que el **lenguaje de transmisión de la información sea diferente al del conocimiento**.

La **información** conlleva una interpretación de los datos y es la diferencia de contextos lo que hace que un mismo hecho sea interpretado de distintas maneras y que las personas aporten diferentes puntos de vista o interpretaciones sobre los hechos objetivos. Sin embargo, cuando se transmite el **conocimiento** no se envían bits o se transmite una interpretación, sino que se establece un intercambio o una transacción entre lo que el emisor y el receptor saben.

Se trata de aprovechar las tácitas y a menudo subjetivas percepciones, intuiciones y corazonadas de las personas.

Se puede decir que el *lenguaje utilizado en la transmisión del conocimiento es más emocional* provocado por la mayor participación del recurso humano, entrando en juego las interpretaciones de cada uno de ellos.

El **conocimiento** es **interno**, sin embargo la **información** es algo **externo** que se halla a nuestra disposición fácilmente en cualquier tipo de soporte.

El **conocimiento** posee esta característica ya que reside en la mente de las personas aunque esto no quiere decir que no vaya a proyectarse hacia el exterior o lo que es lo mismo que no vaya a ser transferido.

Tanto la información como el conocimiento son recursos intangibles en las empresas e-business pero sí que es cierto, que por las características (subjetivo, interno..) y la propia definición de **conocimiento**, es éste un recurso que implica **mayores dificultades de imitabilidad** por parte de las empresas del entorno.

La **información se transfiere fácilmente** ya que se trata de un bien público y al ser externa hace que sea más accesible y por tanto de **fácil imitabilidad** por parte de las empresas del entorno.

Si todo el conocimiento que tiene una empresa fuera información dejaría de ser probablemente el recurso más valioso, ya que *la ventaja competitiva basada únicamente en la información no es tan duradera* debido a que ésta de no protegerla o mantenerla en secreto dentro de los límites de la organización, sería mucho más accesible para el entorno competitivo.

Por consiguiente, el valor **no** reside tanto en la **información** sino en el **conocimiento**, eso sí, la información es el vehículo por el que se intenta evocar una respuesta humana.

Para concluir se va a ilustrar la diferencia entre información y conocimiento siguiendo un ejemplo. En una empresa se están vendiendo mensualmente del producto X 300 unidades y 800 unidades del producto Y. Para transformar esta información en conocimiento es necesario extraer conclusiones y valoraciones que sean interesantes para la empresa. Por tanto, si la empresa está vendiendo más del producto Y que de otros debería centrar sus esfuerzos en la venta de dicho producto dado que parece que es el que más valoran los clientes porque satisface mejor sus necesidades y además la tendencia del mercado va en esa línea.

Lo interesante sería dotar a este producto de **características distintivas** que puedan llamar la atención del cliente, para fidelizarlo y así conseguir la mayor cuota de mercado.

Esto ya no es un dato dotado de significado sino que la empresa va a *tomar una serie de decisiones* a partir de lo que demandan los clientes y de las posibles maniobras que crea que en ese momento pueden llevar a cabo otras empresas del entorno. Esto es **conocimiento**.

### **3 .HERRAMIENTAS INFORMÁTICAS PARA GESTIONAR EL CONOCIMIENTO.**

Las empresas de esta era económica se encuentran con grandes cantidades de datos que representan una mina de informaciones.

Por eso, las empresas han de disponer de las herramientas adecuadas para almacenar los **datos**, posteriormente transformarlos en **información** y ésta en **conocimiento**. Así la arquitectura global sobre la que se van a almacenar los datos se denomina *Datawarehouse*



y sobre éste se van a aplicar diversas *herramientas de ayuda a la decisión* tales como: Datamining, Olap y CRM que van a facilitar el proceso de toma de decisiones en la empresa.

Todas ellas se denominan soluciones ***Business Intelligence*** y además de ayudar en el proceso de toma de decisiones, van a permitir generar aplicaciones de *marketing one to one* dirigidas hacia los clientes, consiguiendo de esta forma una fidelización efectiva de los mismos.

### 3.1 DATAWAREHOUSE.

Es la plataforma que concentra toda la información que es de interés para la organización, constituyendo ésta la materia prima que van a manejar las herramientas anteriormente citadas, entre otras.

La definición clásica del Datawarehouse dada por Bill Inmon en su obra de referencia *using de datawarehouse* es la siguiente :

El Datawarehouse es una colección de datos **orientados al tema, integradas, no volátiles e historiadados**, organizados para el apoyo de un proceso de ayuda a la decisión. Dichas características se van a detallar a continuación:

- a) **Orientadas al tema**, el Datawarehouse se organiza alrededor de los temas principales de la empresa. Así los datos se estructuran *por temas*, contrariamente a los datos de las organizaciones tradicionales organizadas por proceso funcional.
- b) **Datos integrados**, hace referencia al nombre de las variables, formato de los distintos campos y tamaño de los atributos.
- c) **Datos historiadados**, en el Datawarehouse los datos no se actualizan nunca, por tanto, éste almacenará el conjunto de valores que un determinado dato ha tenido a lo largo de su historia.
- d) **Datos no volátiles**, la no volatilidad viene a ser una consecuencia de la historialización descrita anteriormente. No se introduce el valor actual de los datos sino que se introducirá un nuevo registro con el valor actual.

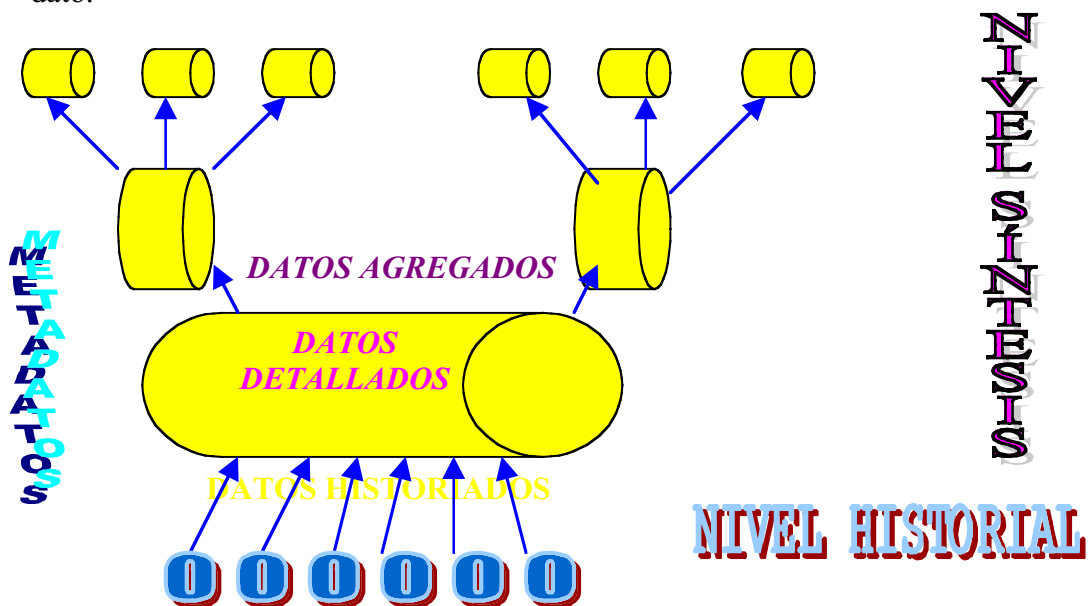
#### *3.1.1 Estructura del Datawarehouse*

El datawarehouse como se ha mencionado, es la plataforma que aglutina datos procedentes del interior y exterior de la empresa, ahora bien, en concreto en el mismo se pueden distinguir *cuatro clases de datos*, organizados según un *eje histórico* y un *eje sintético*. En primer lugar se distinguen los **datos detallados** que son los que proceden de los sistemas operacionales y han sido ya depurados y filtrados.

A continuación se encuentran los **datos agregados** que han sufrido un nuevo proceso de transformación. Esto es, a partir de los *datos detallados* se ha realizado un análisis más profundo intentando sintetizar aquella información más relevante en ese momento según las necesidades del usuario.

Se puede decir que los *datos agregados* son más compactos que los *datos detallados* debido a su *cualidad sintética*. Para la comprensión de los *datos agregados* se hace necesario precisar con más detalle algunas de las características de éstos para así poder realizar un mejor análisis y síntesis de los mismos. Son los **metadatos** los que precisamente van a permitir comprender mejor estos datos ya que su función es enriquecer el dato almacenado en el Datawarehouse. Representan todas las informaciones necesarias para el acceso, comprensión y utilización de los datos del Datawarehouse: semántica, origen, reglas de agregación, formato, etc...

Por último se encuentran los **datos historiados**, y como se ha comentado en una de las características del Datawarehouse, los datos provenientes de los sistemas de producción no destruyen los anteriores valores, sino que crea una nueva ocurrencia del dato.

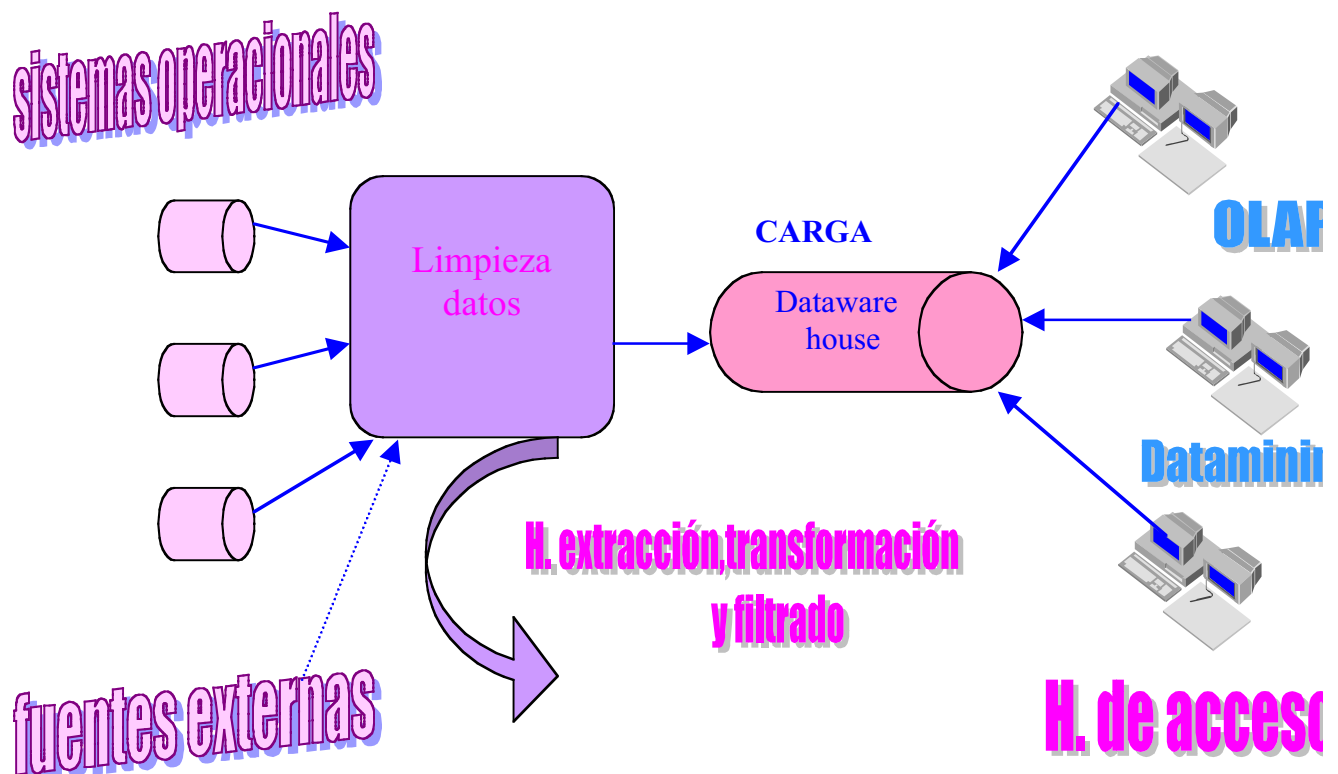


### 3.1.2 Construcción del Datawarehouse

El marco general del Datawarehouse compone tres ámbitos principales: las *aplicaciones*, los *componentes funcionales* y las *infraestructuras*.

Las **aplicaciones**, también llamadas *iniciativas* constituyen cada uno de los temas que se van a tratar en el Datawarehouse, estando el perímetro de cada aplicación claramente definido.

Los **componentes** del Datawarehouse se ven reflejados por medio de este gráfico:



El primer componente son las **fuentes de datos**, esto es, los datos que se van a almacenar en el Datawarehouse proceden de los *sistemas operacionales* de la empresa o del *exterior* de la organización. Los sistemas operacionales mecanizan procesos operativos, muy estructurados y repetitivos, que vienen a constituir las funciones básicas de la empresa, tales como facturación, nóminas etc.. De éstos se puede extraer información válida para las transacciones del día a día pero no sirven para realizar análisis más avanzados.

Por eso, otro de los elementos básicos del Datawarehouse son las **herramientas de extracción, transformación y filtrado**. Éstas se van a encargar de recuperar los datos que

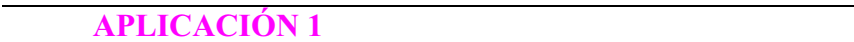
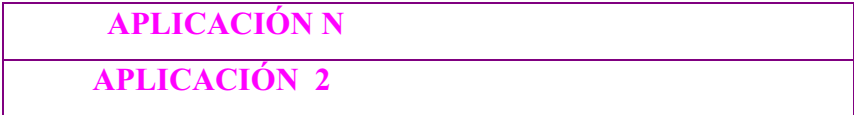
resulten de mayor interés de las aplicaciones transaccionales y los readaptan para su uso informacional, asegurando su calidad e integridad.

Así, el Datawarehouse va a almacenar datos o información depurada y seleccionada teniendo en cuenta las necesidades de la empresa.

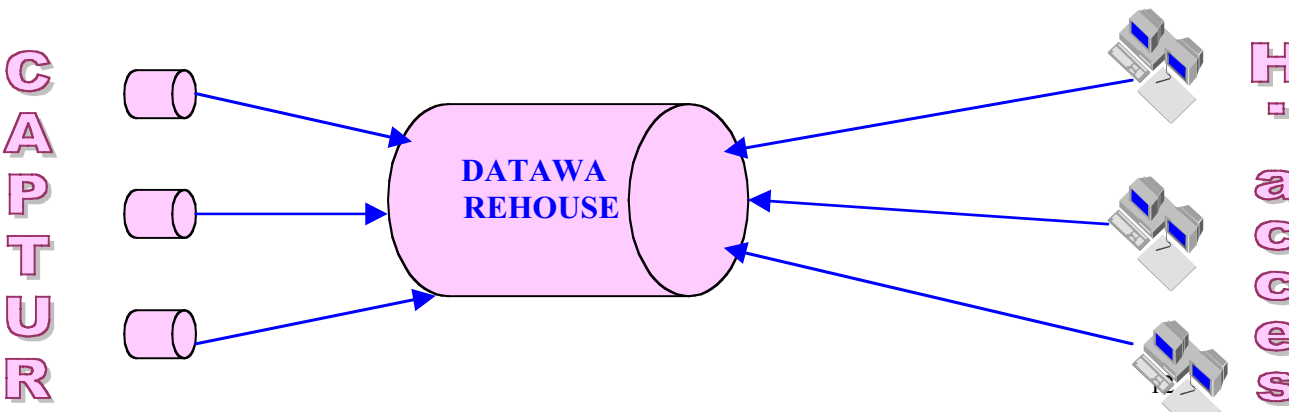
El último componente que actuará sobre el Datawarehouse son las **herramientas de acceso** (Datamining, Olap, etc....), que van a permitir al usuario acceder a dicha *información* para transformarla en *conocimiento*.

Por último, hay que considerar dos niveles de infraestructura en el Datawarehouse: la *infraestructura técnica* y *operativa*.

La *infraestructura técnica* está formada por el conjunto de componentes materiales y programas y la *operativa* lo componen el conjunto de procedimientos y servicios para administrar los datos y utilizar el sistema.



# INICIATIVAS





INFRAESTRUCTURA OPERATIVA
INFRAESTRUCTURA TECNICA

### *3.1.3 Arquitectura del Datawarehouse*

Se pueden distinguir tres tipos de arquitecturas: la arquitectura *real*, la arquitectura *virtual* y la arquitectura *remota*.

La arquitectura *real* es la más utilizada por las empresas de la actualidad. El almacenamiento de los datos del Datawarehouse se realiza en un SGBD separada de los sistemas operacionales. Como se ha mencionado, los datos cargados en el Datawarehouse previamente han sufrido un proceso de limpieza y transformación.

En la arquitectura *virtual* los datos del Datawarehouse residen en el sistema de producción, esto es, no existe una arquitectura real donde depositar una vez filtrados, aquellos datos de interés para la organización.

La arquitectura *remota* es una combinación de los dos tipos de arquitecturas anteriores.

Ahora bien, de nada le sirve a la empresa disponer de una herramienta como el Datawarehouse, sino va a ser capaz de apoyarse en herramientas informáticas que le permitan o al menos faciliten generar conocimiento dentro de la empresa.

Luego el Datawarehouse constituye la arquitectura sobre la que se van a engazar las distintas *herramientas de ayuda a la decisión*, utilizando cada una de ellas diferentes técnicas o procesos para conseguir el objetivo último, esto es, generar **conocimiento**, a partir de los **datos e información** contenida en el Datawarehouse.

### 3.2 DATAMINING O MINERÍA DE DATOS

Se define como la herramienta que permite la extracción de *información oculta* y predecible de grandes bases de datos, es una poderosa tecnología nueva con gran potencial para ayudar a las empresas a concentrarse en la información más importante de sus bases de información (Datawarehouse).

Esto es, va a permitir extraer conocimiento a partir de los datos e informaciones que se encuentran en el Datawarehouse, llegando a crear modelos implícitos en los datos.

El Datamining va a permitir descubrir **conocimiento nuevo**, *información no evidente, desconocida a priori y potencialmente útil* que va a facilitar una nueva visión a los problemas, ya que las consultas que se realizan en el almacén de datos, buscan relaciones entre distintos productos y servicios adquiridos por una misma persona en una empresa, la segmentación de los clientes o el comportamiento de las ventas de un artículo para una mejor y más eficaz campaña de marketing.

Así, la minería de datos va a descubrir **relaciones nuevas** entre distintas variables, bien sea entre productos, distintos tipos de clientes etc...de forma que se obtendrá **nuevo conocimiento** que ni siquiera podía suponerse que se pudiera generar en base a esa combinación de informaciones o datos contenida en la base de datos.

La minería de datos trabaja con **datos individuales, concretos**, descubriendo las regularidades y patrones que presentan entre sí y generalizando a partir de ellos.

Por tanto, el Datamining será capaz de responder a preguntas como ¿qué tipo de clientes tiene la empresa? o ¿cuál es el perfil típico de cada clase de usuario?.

Se dice que la minería de datos realiza un **análisis exploratorio, no corroborativo** porque trata de descubrir *conocimiento nuevo* pero en ningún caso confirma o desmiente hipótesis. Por tanto, no es necesario tener una idea concreta de lo que se está buscando, ya que es el **sistema y no el usuario** el que encuentra las hipótesis, además de comprobar su validez.

Esto no implica que sean las técnicas de minería de datos las que toman las decisiones, obviamente son las *personas* las que con el **conocimiento** adquirido y la *experiencia*, tomarán las decisiones oportunas, tratando de mejorar los resultados de la empresa.

La minería de datos está compuesta por un conjunto de *herramientas* (agrupación, asociación, secuenciamiento, simulación, previsión etc...) que no van a perseguir cada una de ellas los mismos objetivos.

Por eso las diferentes *técnicas* en las que se basa el Datamining también diferirán a la hora de aplicarse en función de la herramienta utilizada.

Una de las *técnicas* más utilizadas son las *estadísticas* y entre éstas la *regresión lineal*, ya que va a permitir establecer *relaciones* entre los datos. Consiste en descubrir relaciones entre variables X e Y del tipo:

$$Y=aX+b, \text{ donde } a \text{ y } b \text{ son dos valores reales.}$$

Esta técnica es rápida y eficaz pero insuficiente para espacios *multidimensionales*. Esto es, la mayoría de las ocasiones las herramientas de Datamining van a extraer **conocimiento nuevo** pero a partir de la relación entre *varias variables* contenidas en la base de datos. Por eso, las herramientas de Datamining utilizarán otro tipo de técnicas más sofisticadas como las *redes neuronales*, *análisis discriminante*, *árboles de decisión*, *aprendizaje no supervisado*, *detección de desviaciones*, *técnicas de visualización etc...* para llevar a cabo el análisis.

La minería de datos es una herramienta muy útil para aplicar en muchos ámbitos pero quizá es en **marketing** donde mayor éxito tiene.

Así, va a permitir detectar entre los potenciales clientes, los que presentan una mayor probabilidad de responder a una determinada campaña publicitaria, detectar nuevas oportunidades de mercado y comparar hábitos de consumo entre diferentes clientes.

Resulta adecuado para la realización de análisis de **predicción**, esto es, a partir de los datos almacenados en el Datawarehouse, es capaz de elaborar modelos que permitan *estimar con precisión* la evolución de una variable en el futuro. Esto puede resultar interesante en la detección de oportunidades, prevención de problemas, gestión óptima del personal y optimización de stocks.

Así mismo, es de utilidad para reducir riesgos, detectar fraudes, controles de calidad etc...

### 3.3 HERRAMIENTAS OLAP (On line Analytical Processing)

Estas herramientas van a permitir realizar un análisis rápido de información *multidimensional*. Van a interpretar una tabla con n atributos independientes como un espacio n-dimensional.

Por ejemplo, un análisis de ventas puede exigir tratar la información global de las ventas por territorio, agentes, periodos, artículos etc, donde como es fácil imaginar, son múltiples las *vistas* que pueden obtenerse de la información.

Ahora bien, no necesariamente el análisis se va a llevar a cabo en bases de datos multidimensionales siempre ya que existen otras arquitecturas dentro de estas herramientas OLAP: ROLAP, MOLAP, HOLAP, donde el análisis se realiza de diferentes maneras.

Así, las herramientas OLAP, en vez de considerar cada transacción individualmente, van a interrelacionar distintas variables llevando a cabo un *análisis agregado de toda la información*. El éxito de esta nueva tecnología es debido en gran parte a su capacidad

gracias a su diseño, para tratar *grandes volúmenes de información no transaccional muy desagregada*, que debe ser analizada bajo muy diferentes perspectivas, denominadas *dimensiones*. Entre éstas el *tiempo* es una de las más frecuentes, ya que se analizarán las diferentes magnitudes del negocio en función de esta variable, obteniéndose una visión clara e interrelacionada de los distintos datos.

El análisis que realizan las herramientas OLAP está *dirigido por el usuario y es deductivo*, esto es, el usuario parte de una hipótesis o pregunta y este tipo de herramientas van a relacionar los distintos datos que estén contenidos en el Datawarehouse para dar respuesta a la misma, obteniendo así conocimiento que puede ser de interés en el proceso de toma de decisiones.

Por eso se dice que el análisis que llevan a cabo estas herramientas es *corroborativo*, esto es, tratará de confirmar o desmentir hipótesis. Por tanto, es necesario tener una idea concreta de lo que se está buscando, ya que **es el usuario y no el sistema** el que lo confirma o desmiente.

Además, las aplicaciones OLAP trabajan generalmente con *datos agregados*, esto es lógico, teniendo en cuenta que el análisis que se realiza es *multidimensional* y se interrelacionan muchas variables que van a permitir obtener una visión global del negocio.

Llegado a este punto, se puede observar como las dos herramientas mencionadas utilizan modos muy distintos de obtener **conocimiento** a partir de los **datos e información** contenidas en las bases de datos, por eso se hace preciso que se complementen en todo momento.

Así, sería interesante incidir de forma más clarificadora en las diferencias existentes entre ambas a la hora de obtener conocimiento.

	<b>DATAMINING</b>	<b>OLAP</b>
Obtención conocimiento	Nuevo	Multidimensional
Análisis	Exploratorio	Corroborativo
Dirigido	Sistema	Usuario
Razonamiento	Inductivo	Deductivo
Trabaja con datos	Individuales / Concretos	Agregados

Por medio de un ejemplo, se puede observar de forma práctica lo expuesto teóricamente:



Una pregunta típica del sistema OLAP sería "El año pasado, ¿se compraron más furgonetas en Madrid o en Barcelona? . La respuesta dada en este caso sería del tipo "En Cataluña se compraron 15000 furgonetas, mientras que durante el mismo intervalo en Madrid, se compraron 12000".

Sin embargo, un problema típico para resolver la minería de datos sería "Hallar un modelo que determine las características más relevantes de las personas que compran furgonetas". A partir de los datos del pasado el sistema de minería de datos proporcionaría una respuesta del tipo: "Depende de la época del año y situación geográfica". "En invierno los habitantes de Madrid que pertenecen a un cierto grupo de edad y nivel de ingresos probablemente comprarán más furgonetas que personas de las mismas características en Cataluña."

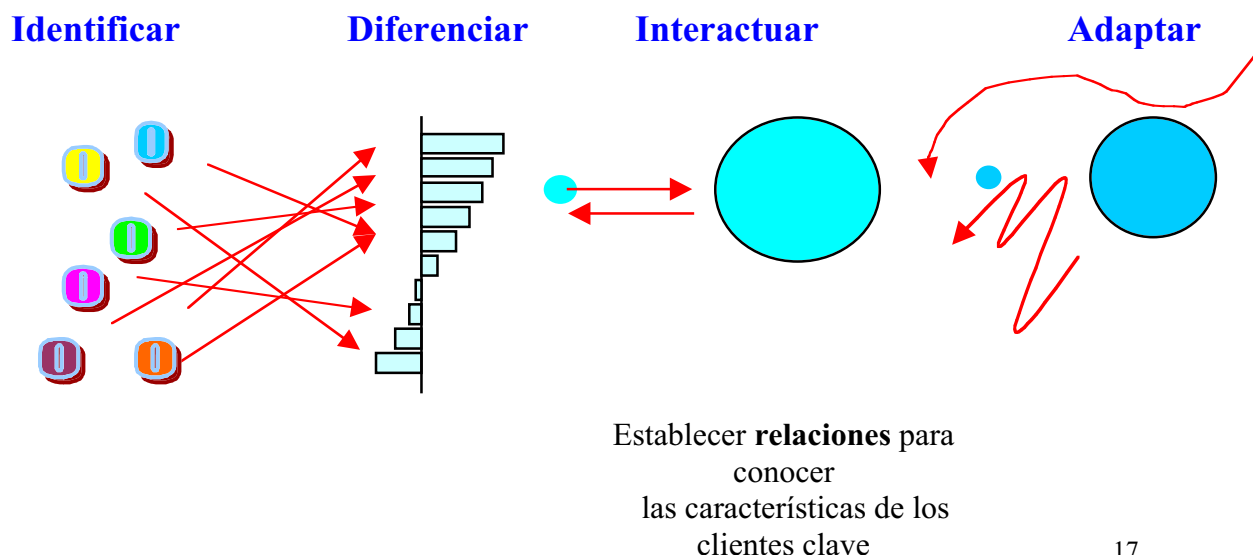
### 3.3 C.R.M (CUSTOMER RELATIONSHIP MANAGEMENT).

El C.R.M es una *estrategia centrada en el cliente*, que busca un crecimiento en beneficios a través de proporcionar un mayor valor al cliente (Mark Roger).

Así como las otras herramientas obtenían **conocimiento** que podía ser útil para tomar decisiones referentes a cualquier ámbito de la empresa, esta técnica va a revelar únicamente la *información relativa a los clientes* con el conocimiento adquirido a partir de la información procedente de las bases de datos.

El objetivo del C.R.M es facilitar el exámen y la comprensión de los factores complejos que determinan la actividad del cliente, con el objeto de ofrecer una *calidad del servicio irreproachable, fidelizarlo y ganar en rentabilidad*.

Para ello, se hace necesario crear una **economía cliente-céntrica** que pivote en su conjunto alrededor del cliente y sus necesidades.



"Seleccionar" a  
cada cliente de manera  
inequívoca

Ordenarlos  
por margen de contribución,  
beneficios que puedan  
reportar etc..

Dar a los clientes  
clave aquello que **satis**  
sus necesidades

Ahora bien, para que la información pueda ser transformada en conocimiento ésta debe ser capturada con inmediatez, seguridad y en tiempo real. Por eso, se hace indispensable disponer de toda la tecnología necesaria para conseguirlo.

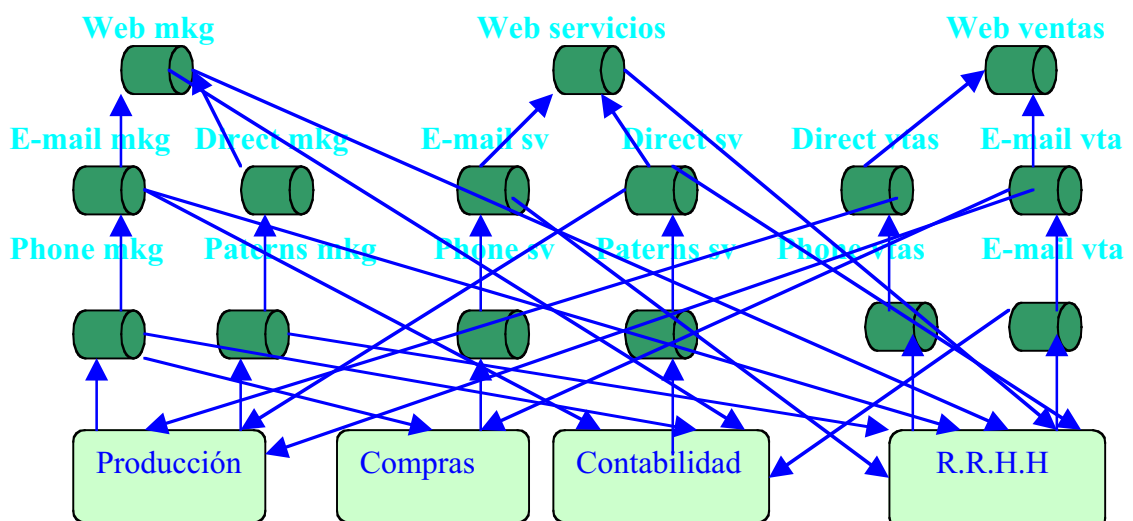
Así el C.R.M dispone de tecnología basada en bases de datos (Datawarehouse y otras) pero fundamentalmente páginas web, centros de atención telefónica, e-mail etc... y todos los medios mediante los cuales una empresa puede interactuar con sus clientes, ya que es el objetivo último del C.R.M.

Pero la empresa no debe centrarse en la Web como único medio de obtener datos e información, pero si que debe *pensar primero en la web*.

Luego la empresa debe tener claro que la información sobre sus clientes se encuentra principalmente en tres áreas de la empresa: *marketing, ventas y servicios*. Ahora bien, no es suficiente con que ésta piense en uno o dos departamentos donde por medio de la web u otros medios, pueda conseguir información acerca de sus clientes.

Es necesario que tenga una visión mucho más global del recurso información y que piense en la empresa en su conjunto.

Esto es, la información que pueda encontrarse en determinados puntos de la empresa, debe ser compartida por el resto de los departamentos de la empresa ya que un dato concreto de un cliente puede ser de gran interés para el resto (producción, contabilidad, compras etc.....) y ayudar al proceso de toma de decisiones.



En las empresas actuales aplicar el C.R.M *es importante* fundamentalmente porque:

- a) Como media, las empresas pierden a la mitad de sus clientes en un periodo de cinco años, la mitad de sus empleados en cuatro años y la mitad de sus inversores en un año.
- b) Descubrir un nuevo cliente es de cuatro a diez veces más caro que retener a un cliente existente.
- b) Reducir los abandonos en un 5% puede aumentar los beneficios hasta un 85%.
- c) El 20% de los clientes de una empresa suelen generar hasta el 80% de los beneficios.

Son razones suficientemente importantes como para plantearse la aplicación del C.R.M en las empresas actuales. Por eso, es importante que la empresa capte en cada momento los **datos e información** relativa a sus clientes para saber qué es lo que demandan y en base a ello satisfacer sus necesidades con productos que les resulten más atractivos que los de la competencia.

Por tanto, la empresa ha de ser consciente de que *no lo sabe todo* y que debe adquirir *conocimiento nuevo constantemente* ya que además las preferencias de los consumidores pueden fluctuar de forma importante.

Si la empresa lleva a cabo este tipo de acciones podrá conseguir el objetivo último que *es retener al cliente*.

#### **4. CONCLUSIONES**

Es importante que las empresas de la *sociedad de la información* utilicen herramientas *Business Intelligence* ya que además de facilitar el proceso de toma de decisiones obteniendo distintos "tipos" de conocimiento, por medio de razonamientos y técnicas diferentes, su aplicación va a reportar una serie de *beneficios* para las empresas. Así, van a permitir, obtener una **visión única** de las necesidades y perfiles de los clientes, mejorar la **productividad** comercial, mejorar la calidad y **satisfacción** de los clientes,

**reducir** riesgos comerciales, **optimizar** las campañas de marketing, **adaptar** la oferta a las necesidades reales del mercado y proporcionar **respuestas rápidas** y completas a las cuestiones de negocio.

### **NOTA BIBLIOGRÁFICA**

**AGUADERO, F (1997):** *La sociedad de la información*. EDITORIAL ACENTO. Madrid

**CORNELLA, A (1994):** *Los recursos de información*. MC GRAW HILL. Bilbao

**CORNELLA, A (2000):** *Infonomía.com: la empresa es información*. EDICIONES DEUSTO. Bilbao

**DANS, E:** *Encuentros sobre CRM.. MARZO 2001*. Bilbao

**ESCOBAR, M (2000):** "La empresa e-business: transformación, modelo de gestión y planificación estratégica", *Economía Industrial, la organización para la innovación(II)*. págs 101-109.

**FRANCO, J. M Y EDS-INSTITUT PROMÉTHÉUS (1997):** *El Datawarehouse, El Datamining*. GESTIÓN 2000. Barcelona

**GARCIA, C, PASCUAL, F, CHAPARRO, J:** *Sistemas de información y comunicación para la gestión del conocimiento*. <http://www.fcee.ulpg.es>.

**GIL PECHUÁN, I (1997):** *Sistemas y Tecnologías de la información para la gestión..* MC GRAW HILL. Madrid

**INMON, W.H, RUDIN, K, BUSS, CK, SOUSA, R (1999):** *Datawarehouse performance*. WILEY COMPUTER PUBLISHING. Canadá

**PERLADO, G (2000):** "Los sistemas CRM, elemento clave en la relación con los clientes", *Economía Industrial, la organización para la innovación(II)*, págs 79-89.