

# Modelado de un Sitio *Web* para Mejorar la Relación con los Usuarios

por

Rodrigo Alfaro Arancibia  
 Ingeniero Civil Industrial de la Universidad Católica de Valparaíso  
 Académico de la Escuela de Ingeniería Industrial de la Universidad Católica de Valparaíso  
 e-mail: rodrigo.alfaro@ucv.cl

## Resumen

El gran auge de Internet, el crecimiento explosivo de transacciones vía *web* y la formalización de filosofías de orientación al cliente, integración de funciones, coordinación con proveedores y vendedores, etc., motiva a que un gran número de empresas intenten realizar ventas en línea a través de la red.

Por otra parte, recientemente se ha desarrollado un área de investigación denominada *Web Mining*, que combina otras dos áreas: *Data Mining* y *World Wide Web*, cuyo objetivo es utilizar la información que se registra en los sitios *web*, generalmente en forma desordenada, para apoyar decisiones de diseño, comerciales, marketing, etc.

No obstante, muchos sitios de comercio electrónico se elaboran y ponen en marcha sin una adecuada planificación y modelado, lo que se traduce en que no se utilicen todas las potencialidades que entrega Internet o las nuevas Tecnologías de la Información. Ello confirma la necesidad de enfatizar que se debe modelar conceptualmente un sistema antes de implementarlo.

En este artículo se describe un método para elaborar un modelo que registre la relación o interacción de los clientes o visitantes con un sitio *web*. De esta forma se puede obtener información para mejorar esta relación en distintos aspectos, tales como: estructura del sitio, promoción de productos a clientes específicos, precios dinámicos, negociación en línea, etc.

Palabras Claves (*Key Words*) : *Web Mining*, *Data Mining*, CRM, Comercio electrónico, Internet.

## 1. Introducción

El gran auge de Internet y el crecimiento explosivo de transacciones vía *web*, motiva a que un gran número de empresas intenten realizar ventas en línea a través de la red. Sin embargo, muchos sitios de comercio electrónico se elaboran y ponen en marcha sin una adecuada planificación y modelamiento, lo que se traduce en que no se utilicen todas las potencialidades que entrega Internet o las nuevas Tecnologías de la Información.

## Abstract

*The great peak of Internet, the explosive growth of transactions via web and the formalization of orientation philosophies to the client, integration of functions, coordination with suppliers and salespersons, etc., it motivates to that a great number of companies tries to carry out on-line sales through the net.*

*On the other hand, a new area of investigation denominated Web Mining that combines other two areas: Data Mining and World Wide Web whose objective is to generally use the information that registers in the places web, in disordered form, to support design decisions, commercial, marketing, etc.*

*Nevertheless, many electronic commerce sites are elaborated and they start without an appropriate planning and modeling, what is translated in that all the potentialities are not used that gives Internet or the new Information Technologies. It confirms it the necessity to emphasize that it should be modeled a system conceptually before implementing it.*

*In this article a method is described to elaborate a model that registers the relationship or the clients' interaction or visitors with a web site. This way one can obtain information to improve this relationship in different aspects, such as: structures, promotion of products to specific clients, dynamic prices, on-line negotiation, etc.*

Por otra parte, las nuevas filosofías de orientación al cliente (*Customer Relationship Management*), integración de funciones (*Enterprise Resource Planning*), coordinación con proveedores (*Supply Chain Management*) y vendedores (*Selling Chain Management*), etc., tienen un potencial de desarrollo mucho mayor debido a la explosiva conectividad y riqueza de la comunicación que se puede establecer por medio de Internet.

Se vuelve a confirmar entonces la necesidad de enfatizar que se debe modelar conceptualmente un sistema antes de implementarlo, describiendo qué hará el sistema, para posteriormente determinar cómo deberá hacerlo.

Este artículo, extractado de la Tesis para optar al grado de Magíster en Ingeniería Industrial, mención Gestión de la Universidad Católica de Valparaíso [1], fundamenta la necesidad de elaborar un modelo que registre la relación o interacción de los clientes o visitantes con un sitio *web* y proponer una metodología para hacerlo. De esta forma, se puede obtener información para mejorar esta relación en distintos aspectos, tales como: estructura del sitio, promoción de productos a clientes específicos, precios dinámicos, negociación en línea, etc.

## 2. ¿Por qué modelar conceptualmente un sitio *web*?

Los modelos se crean para lograr un mejor entendimiento de un ente a representar; si el modelo busca representar un ente que existe, será descriptivo y si el ente no existe, será un modelo prescriptivo.

En fases iniciales del desarrollo de sistemas, específicamente en la fase de Análisis, se deben crear los modelos conceptuales del sistema. Según [4], estos modelos deben describir el problema y los requerimientos de los usuarios, considerando que en una etapa posterior de diseño se construirán los modelos que entreguen la solución que satisfaga los requerimientos de los usuarios y el problema planteado. Durante el análisis, el analista debe poner especial atención en el dominio del conocimiento, objetivos, requerimientos y entorno del sistema, es decir, debe “orientarse al problema” y no limitarse a modelar la solución al problema. De esta forma se facilita el modelamiento y la validación por parte de los usuarios.

Los modelos creados hacen uso de notación gráfica que representa principalmente la información, los procesos y el comportamiento del sistema. Esto puede complementarse con texto, ya sea en lenguaje cotidiano o especializado. El rol de los modelos se puede sintetizar en:

- Ayudar al analista a entender la información, la función y el comportamiento del sistema, haciendo por tanto más fácil y sistemática la tarea de analizar requerimientos.
- Convertirse en el punto de comparación entre lo logrado y lo planificado.
- Fundamentar el diseño, proporcionando al diseñador una representación lógica, o esencial, de la implementación.

Por otro lado, se puede destacar la gran popularidad que han logrado las aplicaciones *web*. Esto debido al rápido progreso de herramientas y tecnologías para desarrollarlas, y a que los diseñadores de sistemas están reconociendo que hay situaciones donde las aplicaciones *web* tienen ventajas significativas con respecto a las aplicaciones tradicionales.

Los ambientes actuales del desarrollo facilitan la producción de aplicaciones *web* simples, lo que tiene el efecto secundario de impulsar a los desarrolladores a que construyan aplicaciones sin el adecuado análisis y diseño. Esto a la larga trae diversos problemas si se piensa que cualquier sistema con complejidad no trivial necesita ser diseñado y ser modelado.

Si se está en presencia del desarrollo de un sitio *web*, se puede fundamentar la necesidad de modelar con dos argumentos complementarios: modelar un sitio *web* para ayudar a manejar la complejidad de un sistema, por una parte, y para aumentar la satisfacción y lealtad de los usuarios, por otra.

## 2.1. Modelar para mejorar la calidad del sistema

En [5] destaca la importancia de modelar un sitio *web* para ayudar a manejar la variedad de un sistema y para que las aplicaciones en *web* no se transformen en aplicaciones muy complejas rápidamente.

Si se considera además, que un sistema se debe representar por distintos modelos consistentes entre sí, cada uno con un propósito específico y con un nivel distinto de abstracción, la necesidad de modelar se hace inevitable.

Un elemento básico de un sitio *web* es la página *web*, por lo que será un elemento que aparecerá modelado en todo sitio *web*. [5] propone utilizar UML para representar una página como un objeto, con las propiedades de tal objeto: fuentes, tablas, texto, etc.; y los *scripts* de la página se representarían como métodos de un objeto página.

Por otra parte [6] argumenta que se deben modelar las aplicaciones en *web*, principalmente debido a que las aplicaciones en *web* están aumentando cada vez más, en cuanto a:

- **Cantidad:** proliferación de intranets, extranets y de *e-commerce*.
- **Complejidad:** Evolución de sitios estáticos a sitios dinámicos y aplicaciones en *web*; proyección de aplicaciones, dimensionando los recursos y plazos de su desarrollo con mayor certeza; evaluación de la viabilidad técnica de un proyecto; análisis de alternativas de arquitecturas, implementación, mantención y adaptaciones; evaluación del proyecto.
- **Importancia:** Los negocios pasan a depender de las aplicaciones en *web*.

Además, modelar una aplicación *web* permite mejorar la calidad del sistema, ya que se detectan los errores y se proponen soluciones antes de implementar el sistema; gestionar el proceso de evaluación del sistema; y apoyar la evolución y el crecimiento de la aplicación, ya que existe documentación y se realiza un mayor esfuerzo para entender y modelar el problema.

## 2.2. Modelar para mejorar la relación con los usuarios

Constantemente aparecen nuevos modelos de negocios que pretenden maximizar las ganancias de las empresas, utilizando para ello nuevos conceptos y enfoques, algunos de los cuales dan buenos resultados y otros no, hasta que el tiempo lo demuestra. Uno de estos modelos de negocios actuales se ha asociado al concepto de CRM aunado al impresionante desarrollo de *h-ternet*.

CRM o *Customer Relationship Management* es un proceso por el cual se enfocan las acciones de una organización hacia la satisfacción y lealtad de los clientes, CRM es un proceso que posibilita reorientar la estrategia desde la visión centrada en el producto hacia una perspectiva centrada en el cliente y sus necesidades, proceso por el cual la empresa organiza la información de sus clientes con el fin de conocerlos y, a partir de tal conocimiento, establecer relaciones rentables y duraderas con aquellos clientes que puedan proporcionar una mayor rentabilidad a la empresa.

Las empresas que tradicionalmente han enfocado sus esfuerzos en productos, ahora reconocen la importancia del cliente como el elemento más importante en las relaciones comerciales.

Con el auge de Internet y las facilidades que ésta presenta, se han creado empresas que desde sus inicios se enfocan en las necesidades de los clientes.

Según distintos autores [2], el proceso de CRM consta de tres etapas principales. Estas etapas se describen a continuación:

1. **Integración:** dada la existencia de una serie de fuentes, canales de información y departamentos dentro de las organizaciones, el primer paso esta en la integración de los datos producto de la relación con los clientes. Esta información debiera estructurarse en un *Data Warehouse*. Esta etapa permitirá proceder con el proceso de análisis y extracción de conocimiento del negocio.
2. **Análisis:** esta etapa busca, a partir de los datos de los clientes, alcanzar un conocimiento de las tendencias y patrones de comportamiento que permita crear un modelo de predicción del comportamiento futuro y establecer indicadores que ayuden al soporte de la toma de decisiones. Puede asociarse a un *Data Mining*.
3. **Acción:** esta etapa es la que da sentido a una estrategia de CRM, pues tomando en cuenta los resultados del análisis, es necesario tomar acciones concretas que pueden afectar desde las estrategias de marketing hasta la organización propia de la empresa.

Como se ha mencionado, el objetivo que busca el proceso de CRM es maximizar los beneficios de la empresa mediante un conocimiento de los clientes y la relación que se tiene con ellos.

Ninguna técnica analítica por sí sola puede entregar una vista integral de los clientes, de su comportamiento y de sus características descriptivas. La rentabilidad de los clientes se debe estimar mediante un análisis global que describa a los clientes. Esto incluye: segmentar, medir la rentabilidad y la fidelidad del cliente, identificar características comunes en un segmento de clientes, analizar la respuesta esperada a una campaña de marketing, etc.

Puede considerarse a Internet como un medio de comunicación para llegar a los clientes, pero con la ventaja de que no sólo comunica al igual que lo hace un periódico o la televisión, sino que permite interactuar y transferir información en dos vías, tanto desde la empresa a los clientes como desde los clientes a la empresa; esto da lugar a la interacción. Además, Internet aprovecha la funcionalidad de sus herramientas como el e-mail y el chat, para llegar a establecer una relación con los usuarios. Todo esto es ideal para un proceso de CRM, pues Internet le brinda la facilidad de obtener los datos de los clientes de una forma muy ágil, por ejemplo, este es el caso de aquellas páginas electrónicas que solicitan cierta cantidad de datos para tener acceso a la información contenida en una página *web*.

Las interacciones con los clientes que ocurren a través de Internet aplican idealmente con el proceso de CRM, ya que permiten la utilización de una plataforma tecnológica para el diseño, desarrollo e implementación de técnicas de administración de información como el *datawarehouse* y el *datamining*, los cuales permiten pasar fácilmente de la etapa de integración, mencionada antes, a la etapa de análisis; asimismo, estas técnicas se pueden combinar con herramientas de *e-business* para asociar las interacciones a través de Internet con el valor económico actual y futuro de los clientes.

Otro punto a señalar es que a través de Internet se logra uno de los objetivos fundamentales de CRM, como es la personalización de los servicios o productos. Dentro de la *web* pueden encontrarse muchos ejemplos donde el consumidor puede dar las especificaciones de los productos y servicios de acuerdo con sus gustos y preferencias, antes de adquirirlos.

Se puede decir entonces, que al decidir que se va a desarrollar un sitio de comercio electrónico, hay que partir del supuesto que a los clientes se les debe atender de la mejor manera posible, porque sino, la probabilidad de tener éxito es muy escasa.

### 3. Modelado

En [3] se presenta la notación de la metodología *Object-Oriented Hypermedia Design Method* (OOHDM), que modela conceptualmente un sitio *web* utilizando *Unified Modeling Language* (UML) y propone un modelo complementario para representar el modelamiento navegacional, que se basa principalmente en el respectivo Modelo Estático. Debido a que UML no trata el modelamiento navegacional, se desarrolla la notación del esquema navegacional basándose en UML. A continuación, en la Figura 1, se presenta el Diagrama Entidad Relacionamiento que representa la estructura operacional de un sitio *web* de comercio electrónico.

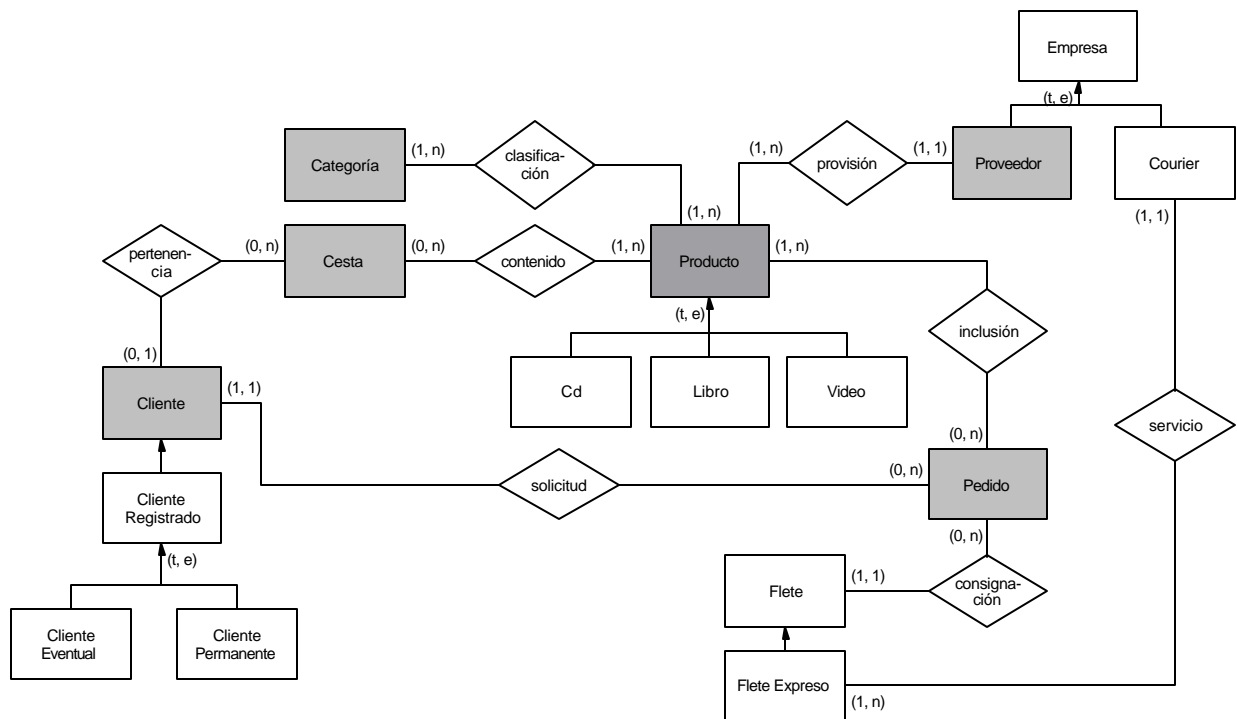


Figura 1 – Diagrama Entidad Relaciónamiento Operacional

De acuerdo a los ejemplos analizados en [7], [3] y otros, se puede inferir que generalmente las navegaciones que debe realizar un visitante para encontrar una instancia de una determinada Clase o Entidad se pueden clasificar en tres grandes tipos. A partir del Diagrama presentado en la Figura 1, se pueden determinar algunos criterios de búsqueda que el sitio debiera ofrecer, tales como:

- **Búsqueda directa del producto requerido:** Este tipo de búsqueda asume que el visitante conoce el sitio y sabe lo que quiere comprar.
- **Búsqueda a través de productos relacionados:** Esta búsqueda la realizarían visitantes que conocen o no el sitio, pero que no saben bien lo que desean comprar, es decir, tienen una idea del producto que desean, o bien, del tema (o concepto) que les interesa, pero no tienen claro que es lo que el sitio les puede ofrecer para sus requerimientos.
- **Búsqueda personal:** Este tipo de búsqueda supone que el visitante ha realizado búsquedas o compras anteriores en el sitio y que se ha registrado, por lo tanto puede buscar lo

que dejó en su cesta de compras, o productos relacionados con sus compras anteriores, etc.

Además, se puede agregar que la alternativa que escoja el visitante para buscar una determinada instancia de alguna clase o entidad, dependerá de su grado de conocimiento en la utilización de aplicaciones en *web*, conocimiento del sitio, conocimiento de lo que se quiere, tipo de compra que desea realizar, etc.

Si se consideran los tipos de búsqueda mencionado anteriormente, se puede realizar una primera categorización de las Entidades que debieran asociarse al esquema navegacional, éstas se muestran ennegrecidas en la Figura 1.

A partir de los criterios de búsqueda basados en el DER Operacional, se puede obtener un DER que registre las interacciones de los usuarios en el sitio *web*. De esta forma, la Figura 2 representa un modelo que registra las interacciones del visitante con el sitio automáticamente, es decir, sin que el visitante tenga que completar encuestas o responder preguntas.

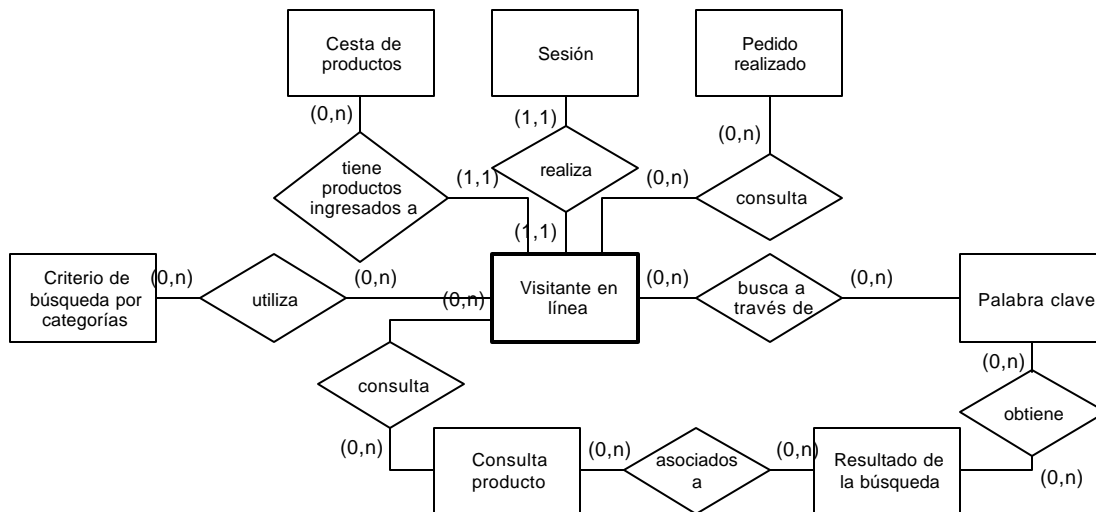


Figura 2 – Diagrama Entidad Relaciónamiento Navegacional

Si se observa la Figura 2, todas las relaciones de la Entidad Visitante en Línea son tomadas a partir de las posibilidades de búsqueda que se presentan en el DER Navegacional de la Figura 1, excepto la Entidad denominada Sesión, ya que esta Entidad es una característica asociada al visitante, que se registra porque puede servir también para mejorar las relaciones con ellos, por ejemplo: registrando el idioma que se utiliza desde ciertas conexiones, la hora local, hora de inicio de la sesión, hora de término, etc.

Una pregunta que nace a continuación es cómo se relacionan, si es que existe relación, el DER Operacional mostrado en la Figura 1 y el DER Navegacional mostrado en la Figura 2.

Rápidamente se pueden vislumbrar asociaciones o relaciones entre algunas Entidades de un modelo con Entidades del otro. A continuación se analiza qué tipo de asociaciones pueden darse entre las Entidades de DER Operacional y del DER Navegacional. Para esto es conveniente describir brevemente los elementos y relaciones que se deben modelar en el DER.

El DER describe cómo se relacionan los elementos estáticos. Este modelo se compone de tres piezas claves: las entidades, los atributos que describen a las entidades, y los relacionamientos que conectan las entidades entre sí. Una entidad es la representación de cualquier composición de información que deba almacenar el sistema, generalmente tiene un gran número de pro-

iedades o atributos diferentes. Una entidad, a diferencia de un objeto, sólo encapsula datos. Además, en un Diagrama Entidad Relaciónamiento pueden existir tres tipos de asociaciones entre Entidades, éstas son:

- **Relacionamiento:** Las entidades se conectan entre sí a través de los relacionamientos. Estos representan las conexiones entre entidades que son relevantes para el sistema. Cada trío “entidad – relacionamiento – entidad” tiene asociada una cardinalidad, ésta representa el número mínimo y máximo de ocurrencias de una entidad (instancia) que se pueden relacionar con ocurrencias de la otra entidad. Generalmente la cardinalidad se expresa como (0,n), (1,1) o (1,n). Se representan como rombos en el DER.
- **Jerarquía de Tipos:** Además, en el Diagrama Entidad Relaciónamiento se pueden definir Jerarquías de Herencia o Tipos, lo que se realiza a través de relacionamientos de subconjunto, donde el conjunto del cual son tomados los subconjuntos es denominado padre o súperentidad y los subconjuntos son denominados hijos o subentidades. Todas las propiedades (o atributos) y relacionamientos del padre son válidos para sus hijos. Un tipo de entidad puede estar involucrado en más de una jerarquía de herencia.
- **Asociaciones en una Entidad Agregada:** En ocasiones es necesario conectar relacionamientos -que ya conectan a dos entidades- con otro relacionamiento, como esto no se puede realizar (relacionar dos relacionamientos) se agrega el conjunto de entidades y relacionamientos, obteniendo una Entidad Agregada, la que tiene todas las características de una Entidad normal, y teniendo en consideración la consistencia de los atributos e identificadores de las entidades que agrega.

Realizando una análisis caso a caso de cada una de las asociaciones se puede obtener el DER que se presenta en la Figura 3.

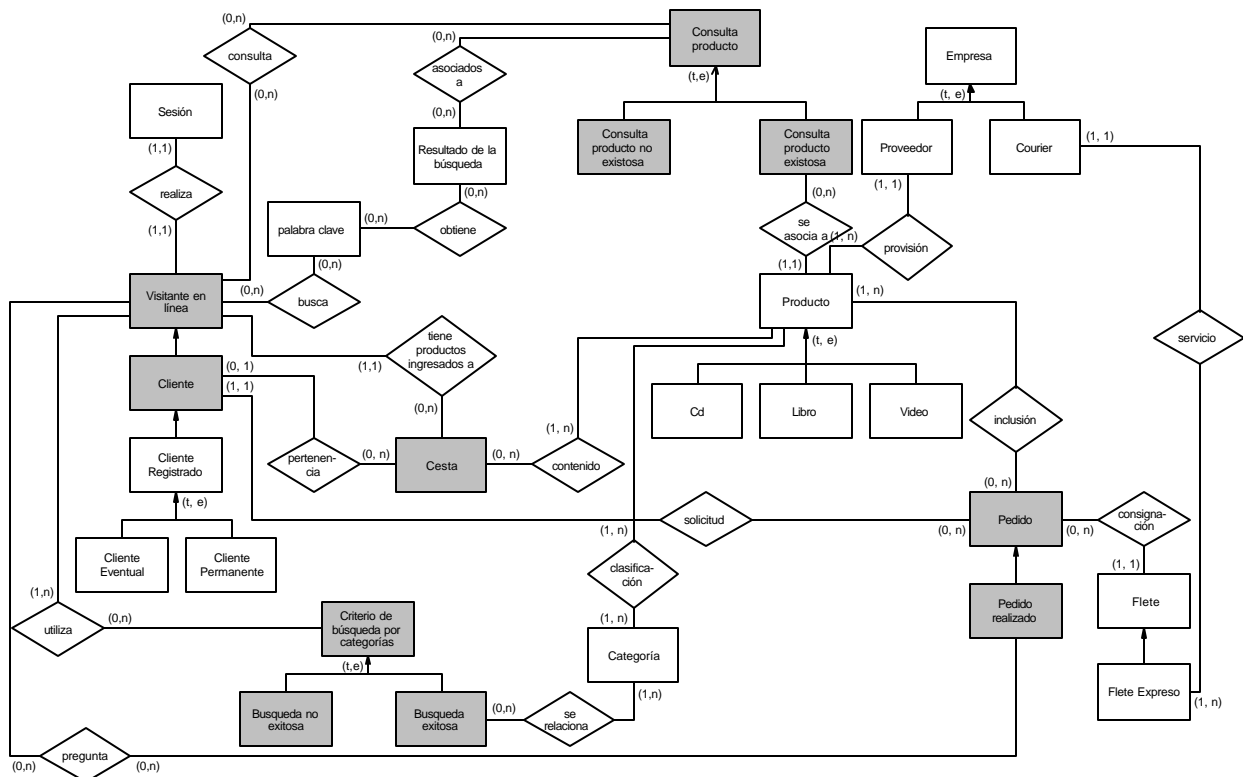


Figura 3 – Diagrama Entidad Relaciónamiento Integrado

## 5. Usos

Dado que el objetivo es mejorar las relaciones en base a los datos registrados, se propone elaborar un *Data Warehouse* o *Data Mart* que contenga datos de la base de datos operacional y datos registrados de las interacciones con los usuarios. Este *Data Warehouse* debiera elaborarse a partir del Diagrama Entidad Relacionamiento Integrado, presentado en la Figura 3. En él se deberán organizar y almacenar los datos requeridos para el procesamiento analítico de la información a largo plazo.

Para diseñar un *Data Warehouse* se deben realizar dos tareas. Primero debe desarrollarse un diseño lógico y a continuación se realiza un diseño físico.

Como el formato de los datos no es relevante para el análisis, los datos de las tablas normalizadas de la base de datos operacional, se deben desnormalizar y en ocasiones, cambiar el formato. Se trata de armar una base de datos informacional a partir de la base de datos operacional.

Además, se deben eliminar las inconsistencias y redundancias que pueden derivarse de la integración de las bases de datos operacionales de distintas áreas de la organización y filtrar los datos muy específicos que no aportarán en las decisiones que el *Data Warehouse* apoyará.

### 5.1 Web Mining

*Web Mining* es la aplicación de técnicas de *Data Mining* para descubrir patrones de los datos de la *web* con el propósito de entender y mejorar el servicio a las necesidades de aplicaciones basadas en *web*. Aunque aún no existe un completo consenso sobre lo que es *Web Mining*, la mayoría de los estudios, tales como [8] y [9] diferencia entre tres áreas: *Web Mining* de contenido, *Web Mining* de estructura, y *Web Mining* de uso.

El *Web Mining* de contenido se enfoca en el descubrimiento y recuperación de información útil de la *Web*, en cuanto a contenido, datos y documentos. Por su parte, *Web Mining* de estructura enfatiza en el descubrimiento de cómo modelar las estructuras de vínculos de la *web*. La distinción entre estas dos categorías no es muy clara a veces. *Web Mining* de uso es una categoría relativamente independiente de las anteriores, pero no aislada. Esta categoría describe, principalmente, las técnicas que descubren patrones de uso de los usuarios de sitios *web* e intenta predecir sus conductas. En [9], propone que el *Web Mining* de uso sea analizado en tres fases distintivas, que son: pre-procesamiento, descubrimiento de patrones y análisis de patrones.

La fase de pre-procesamiento del *Web Mining* de uso se puede asociar a un modelo para registrar los datos de los visitantes en forma ordenada y estructurada. A continuación, en la etapa de descubrimiento de patrones, se utilizaría el *Data Mining*. Finalmente, en la etapa de análisis de patrones se deben tomar las salidas de las etapas anteriores y tomar las decisiones para mejorar las relaciones con los clientes.

Del punto de vista de la técnica, *Web Mining* de uso es la aplicación de técnicas de *Data Mining* a los *logs* (datos de *Web* secundarios) de grandes repositorios de datos *web*. El propósito es producir salidas que puedan usarse en las tareas de diseño, tales como: diseño del sitio *web*, diseño del servidor de *Web* y diseño de la navegabilidad de un sitio *web* [10].

El Pre-procesamiento de contenido es el proceso de convertir texto, imagen, escrituras y otros archivos en las formas que pueden ser usadas por el *Data Mining* de uso.



Para determinar los pre-procesamiento de estructura, se debe considerar que la estructura de un sitio de *web* es formada por *hyperlinks* entre vistas de páginas. De esta forma, los pre-procesamientos de la estructura pueden tratarse en forma similar como el pre-procesamiento de contenido.

Las entradas de la fase del pre-procesamiento pueden incluir *logs* del servidor de *web*, *logs* de destinatarios, archivos de registro, *index sever logs*, y opcionalmente las estadísticas de análisis anteriores.

Las salidas son el archivo de sesión del usuario, archivo de la transacción, arquitectura del sitio, y clasificaciones de la página.

Por su parte, según [9], el pre-procesamiento del uso es, probablemente, la tarea más difícil en el *Web Mining* de uso debido a la falta de datos disponibles. De aquí que la propuesta de esta investigación, en el sentido de desarrollar un DER Navegacional e integrarlo con el DER Operacional, para registrar datos, es de mucho valor.

En el descubrimiento de los modelos y patrones convergen los algoritmos y técnicas de varias áreas de investigación, como *Data Mining*, *Machine Learning*, estadística, y reconocimiento de patrones. De acuerdo a las técnicas utilizadas en estas áreas, se pueden mencionar como las principales: Análisis estadístico, Reglas de asociación, Clustering, Clasificación, Patrón secuencial, Modelamiento dependiente, Redes Neuronales, etc.

#### **4. Conclusiones**

El gran auge de Internet y el crecimiento explosivo de las transacciones vía *web* motiva a que un gran número de empresas realicen ventas en línea a través de la red y aprovechando las particularidades de Internet, adopten una filosofía orientada al cliente, a través de CRM.

Sin embargo, cuantiosos sitios de comercio electrónico se desarrollan sin una adecuada planificación y modelamiento, lo que se traduce en que no se utilicen todas las potencialidades que entrega Internet o las nuevas tecnologías de información.

Este artículo describe la investigación respecto a modelar la relación o interacción de los clientes o visitantes con un sitio *web* de comercio electrónico. La motivación de obtener un modelo es que éste sirva para registrar las interacciones de los visitantes sin tener que consultarles ni molestarles. Se podría decir que sirve para seguir los pasos de los visitantes en el sitio, para conocerlos de mejor forma y poder atenderlos de acuerdo a los gustos, requerimientos y particularidades de cada uno.

Se utiliza un esquema navegacional que representa las interacciones de los visitantes del sitio *web* a través de una metodología propuesta en [3]. Un aporte específico de la investigación es la integración que se realiza entre el DER que registra interacciones del sitio con el DER operacional del negocio.

Se debe hacer notar también, que la integración entre estos diagramas se realiza de forma natural y sin recurrir a otros mecanismos más que los existentes en el modelamiento conceptual del Diagrama Entidad Relacionamiento. En general, se producen relaciones de subconjunto entre algunas entidades del DER Operacional y del DER Navegacional, por lo que se deben crear algunas jerarquías de tipos, las que en definitiva integrarán los dos diagramas obteniendo el Diagrama Entidad Relacionamiento Integrado.

Además, se presenta un área nueva dentro del *Data Mining* como lo es *Web Mining*, enfocándose principalmente en el *Web Mining* de uso. Esta área o disciplina formaliza la necesidad de obtener información de la *web*, y propone metodologías para hacerlo. Sin embargo, esta es un área relativamente nueva, por lo que muchas proposiciones están aún en elaboración.

Cabe destacar que el Diagrama Entidad Relacionamento Integrado es de gran utilidad en la etapa de Pre-procesamiento del *Web Mining* de uso. El Pre-procesamiento es el proceso de convertir texto, imagen, escrituras y otros datos en formas que puedan ser usadas por el *Data Mining* de uso. Las salidas de esta etapa servirán para ser analizadas en la siguiente etapa de Descubrimiento de Patrones. Según [9], el Pre-procesamiento del uso es la tarea más difícil en el *Web Mining* de uso. De aquí que la propuesta de esta investigación es de mucho valor al hacer un Pre-procesamiento más claro y perdurable.

Se propone basarse en la filosofía CRM para obtener información. Sin embargo, aunque esta filosofía sirvió de marco para obtener un tipo de información, no debe ser una restricción para obtener otra información que pueda apoyar a los tomadores de decisiones y perfectamente podría obtenerse información bajo otra filosofía.

## Referencias

- [1] Alfaro, Rodrigo. **Modelado de un Sitio Web de Comercio Electrónico para Mejorar la Interacción bajo la Filosofía CRM**. Tesis para optar al Grado de Magíster en Ingeniería Industrial, Mención Gestión de la Universidad Católica de Valparaíso, Valparaíso, Septiembre de 2001.
- [2] Bryan Bergeron. **Essentials of CRM: A Guide to Customer Relationship Management**, John Wiley & Sons. USA, February 2002.
- [3] Schwabe, Daniel & Vilain, Patrícia. “**Notação da Metodologia OOHDM**”. Río de Janeiro, Abril de 1999. Disponible vía *web* en <http://sol.info.unlp.edu.ar/notacaoOOHDM>.
- [4] Høydalsvik, Geir & Sindre, Guttorm. On the Purpose of Object-Oriented Analysis. **ACM SIGPLAN Notices**, New York, v.28, n.10, Octubre 1993.
- [5] Conallen, Jim. **Building Web Applications with UML**. 4<sup>a</sup> Edición, USA: Addison-Wesley, Junio 2000.
- [6] Deboni, José Eduardo. **Modelando a Web com a UML**. Objetos Distribuídos OD'99. Centro de Convenções Rebouças. Sao Paulo, Brasil, Diciembre 1999. Diapositivas.
- [7] Rossi, Gustavo; Schwabe, Daniel & Lyardet, Fernando. **Abstraction an Reuse Mechanisms in Web Application Models**. ER 2000 Workshops on Conceptual Modeling Approaches for E-Business and The World Wide Web and Conceptual Modeling. Salt Lake City, UTA, USA, October 2000. Editorial Springer-Verlag Berlin Heidelberg 2000, Alemania. Disponible vía *web* en <http://www.telemidia.puc-rio.br/ooohdm/ooohdm.html>.
- [8] Wang, Y. **Web Mining and Knowledge Discovery of Usage Patterns**. CS 748T Project (Part I). February, 2000. Disponible vía *web* en <http://citeseer.nj.nec.com/>
- [9] Cooley, R. **Web Usage Mining: Discovery and Application of Interesting Patterns from web data**. PhD thesis, Dept. of Computer Science, University of Minnesota, May 2000.
- [10] Cooley, R., Mobasher, B., and Srivastava, J. **Data preparation for mining World Wide Web browsing patterns**. Knowledge and Information Systems, 1(1), 1999. Disponible vía *web* en <http://citeseer.nj.nec.com/>