

Diseño de un Almacén de Datos Histórico en el Marco del Desarrollo de Software Dirigido por Modelos

Carlos Gerardo Neil

Directora: Dra. Claudia Pons
Codirector: Dr. Gustavo Rossi

Tesis presentada para obtener el
grado de Doctor en Ciencias Informáticas

Facultad de Informática
Universidad Nacional de La Plata
19 de Noviembre de 2010

1. Introducción

Un *Decision Support System* (**DSS**) asiste a los usuarios en el proceso de análisis de datos en una organización con el propósito de producir información que les permita tomar mejores decisiones. Los analistas que utilizan el **DSS** están más interesados en identificar tendencias que en buscar algún registro individual en forma aislada [HRU96]. Con ese propósito, los datos de las diferentes transacciones se almacenan y consolidan en una base de datos central denominada *Data Warehouse* (**DW**); los analistas utilizan esas estructuras de datos para extraer información de sus negocios que les permita tomar mejores decisiones [GHRU97].

Basándose en el esquema de datos fuente y en los requisitos de información de la organización, el objetivo del diseñador de un **DSS** es sintetizar esos datos para reducirlos a un formato que le permita, al usuario de la aplicación, utilizarlos en el análisis del comportamiento de la empresa.

Dos tipos diferentes (pero relacionados) de actividades están presentes: el diseño de las estructuras de almacenamiento y la creación de consultas sobre esas estructuras. La primera tarea se desarrolla en el ámbito de los diseñadores de aplicaciones informáticas; la segunda, en la esfera de los usuarios finales. Ambas actividades, normalmente, se realizan con escasa asistencia de herramientas automatizadas.

1.1. Problemas y Soluciones

A partir de lo expresado anteriormente Identificamos, por consiguiente, tres problemas a resolver: a) la creación de estructuras de almacenamiento eficientes para la toma de decisión, b) la simplificación en la obtención de la información sobre esas estructuras para el usuario final y, c) la automatización, tanto del proceso de diseño de las estructuras de almacenamiento, como en la elaboración iterativa de consultas por parte del usuario de la aplicación.

La solución propuesta es el diseño de una nueva estructura de almacenamiento que denominaremos *Historical Data Warehouse* (**HDW**) que combina, en un modelo integrado, un *Historical Data Base* (**HDB**) y un **DW**; el diseño de una interface gráfica, derivada del **HDW**, que permite realizar consultas en forma automática y, por último, el desarrollo de un método de diseño que engloba ambas propuestas en el marco del *Model Driven Software Development* (**MDD**).

1.1.1. Creación de Estructuras de Almacenamiento Eficientes

El **DW** es una copia de los datos de las transacciones de una organización, estructurados específicamente, para realizar consultas y análisis [Kim96]; este tipo de almacenamiento juega un rol central en los actuales **DSS** debido a que brindan información crucial para el proceso de toma de decisión estratégica [Inm02].

Una característica distintiva del **DW** es que el tiempo es una de las dimensiones para el análisis [CD97], [GMR98a], pero éste hace referencia al instante en que se realizó una transacción, por lo tanto, no especifica ni cómo ni cuándo han variado, a través del paso del tiempo, los valores de las entidades, atributos e interrelaciones vinculadas a esas transacciones. Si bien el *Temporal Data Warehouse (TDW)* contempla, además de la dimensión temporal, otros aspectos vinculados con el tiempo [HVM99], [EC00], [EKM01], este modelo considera solo las modificaciones que se producen en el esquema del **DW**, tanto en las dimensiones como en las jerarquías.

Por lo tanto, un problema a resolver en este tipo de estructura *Multidimensional (MD)*, en vistas a la necesidad de registrar valores que permitan evaluar tendencias, variaciones, máximos y mínimos, es de qué manera plasmar en el diseño de la estructura **MD** cómo los valores de las entidades, atributos o interrelaciones pueden variar en el tiempo; ya que, aunque los datos necesarios estuvieran almacenados, los mecanismos de búsqueda temporales resultarían complejos [NA02].

1.1.2. Simplificación en la Obtención de la Información

Las herramientas de consulta que dependen de la habilidad de los programadores para un uso eficaz y eficiente, imponen una carga cognitiva que puede disminuir la productividad de los usuarios [ON01]. Resulta, por lo tanto, un desafío para usuarios no técnicos especificar consultas en una estructura de **DB** [LJ09], más aun si los datos están almacenados en una estructura *Temporal Multidimensional (TMD)*.

La forma tradicional de acceder a una base de datos ha sido mediante consultas por medio del *Structured Query Language (SQL)*, un lenguaje diseñado específicamente para crear, organizar y consultar base de datos.

Debido a la complejidad en la formulación de consultas no triviales se han propuesto diversos enfoques para hacerlas más accesible a un espectro mayor de usuarios [FKSS06]; siguiendo esta línea, el uso de lenguajes gráficos para la realización de consultas, comparado con la escritura de expresiones algebraicas, facilitaría sobremanera las tareas del usuario final. Por lo tanto, un lenguaje gráfico debería operar sobre una vista gráfica explícita del esquema conceptual y las consultas deberían ser expresadas sobre la representación gráfica en forma incremental [RTTZ08].

1.1.3. Automatización del Proceso de Desarrollo de Software

Aunque han sido propuestos diversos métodos que permiten derivar el esquema conceptual **MD** a partir de los datos fuentes de la organización y/o de los requerimientos del usuario (ver [CT98], [GMR98a], [TBC99]), la mayoría de ellos deben ser realizados manualmente [RA08]. Por otro lado, un ambiente visual centrado en el usuario debería incluir herramientas automatizadas para la formulación de consultas y proveer diferentes metáforas de visualización [KG95].

Una solución a estos problemas lo plantea **MDD**, este enfoque se ha convertido en un nuevo paradigma de desarrollo de software que promete mejoras en la construcción de software basándose en un proceso guiado por modelos y soportado por potentes herramientas. En **MDD**, la construcción de software se realiza a través de un proceso guiado por modelos y soportado por potentes herramientas que generan código a partir de aquellos. Este nuevo paradigma tiene como

objetivos mejorar la productividad y la calidad del software generado mediante una reducción del salto semántico entre el dominio del problema y de la solución [PGP09].

La idea clave subyacente es que, si se trabaja con modelos, se obtendrán importantes beneficios tanto en productividad, portabilidad, interoperatividad y mantenimiento y documentación [KWB03].

1.2. Nuestra Propuesta: El Data Warehouse Histórico

Por lo considerado anteriormente, respecto al almacenamiento de datos, la recuperación de información y la automatización de ambos procesos, surge la necesidad de un método, en un ambiente asistido por herramientas automatizadas, que faciliten el proceso de diseño de una nueva estructura de almacenamiento que denominaremos **HDW**, propuesta que combina, en un modelo integrado, un **HDB** y un **DW** y cuyo objetivo es resolver las limitaciones temporales de las estructuras **MD** tradicionales y que, además, le permita, al usuario final, realizar sobre esa estructura **TMD**, en forma automática y asistido por una intuitiva interface gráfica, consultas temporales, de toma de decisión o combinaciones de ambas.

1.3. Objetivo de la Tesis

Los objetivos planteados en la tesis son: a) presentar una solución a las necesidades temporales explícitas en los modelos **MD** mediante el diseño de un **HDW**, b) enmarcado en el enfoque **MDD**, derivar una implementación física en un *Relational Data Base Management System (RDBMS)* del **HDW**, c) proponer el diseño de una interface gráfica, derivada del **HDW**, en el marco de **MDD**, para la recuperación de información y, d) resolver en forma automática, mediante sentencias **SQL**, las consultas realizadas por el usuario final en el **HDW**.

La propuesta está resumida en los siguientes ítems:

- La creación de un modelo y un método para el diseño automático de un **HDW** que incluya, además del hecho principal de análisis, estructuras temporales vinculadas a los niveles de las jerarquías dimensionales que posibiliten registrar los datos y recuperar la información que varíase en el tiempo.
- La creación de un lenguaje visual de consultas, derivado de la estructura **TMD** que permita realizar, en forma automática, las consultas temporales y **MD**.
- La aplicación del paradigma **MDD** en la construcción de un **HDW** y de una interface gráfica, que facilite el diseño mediante un proceso guiado por modelos y asistido por herramientas automatizadas que generen código a partir de aquellos, reduciendo, de este modo, el salto semántico entre el dominio del problema y la implementación de su solución.

1.4. Contribuciones Principales

Las contribuciones principales de la tesis son:

- La propuesta de un nuevo modelo de datos temporal (**HDB**) simplificado que permite registrar la variación de los valores de atributos, entidades e interrelaciones que se modifiquen en el tiempo.

- La propuesta de una nueva estructura de almacenamiento de datos (**HDW**), que combina e integra en un solo modelo, un **DW** y un **HDB**.
- La creación de un método de diseño que, a partir de un modelo de datos conceptual y mediante sucesivas transformaciones, permite obtener una implementación lógica de un **HDW** en un **RDBMS**.
- La implementación, mediante el enfoque **MDD**, del método de diseño de un **HDW** mediante la transformación automática de sus modelos abstractos a modelos concretos.
- El desarrollo de un entorno gráfico derivado automáticamente del **HDW**, en el marco **MDD**, para la realización de consultas sobre la estructura **TMD**.
- La generación automática, utilizando el enfoque **MDD**, de sentencias **SQL** que permite realizar, sobre el **HDW**, tanto las consultas características de un **DW** como las típicas de un **HDB**.
- La creación de un prototipo, basada en tecnología **ECLIPSE**, que implemente el método de diseño del **HDW**, la interface gráfica de consultas y la realización de sentencias **SQL**.

1.5. Trabajos Relacionados

Detallaremos, a continuación, una caracterización de los trabajos relacionados vinculados al enfoque de desarrollo utilizado, al diseño de las estructuras de almacenamiento y a las consultas gráficas; luego, estableceremos las principales diferencias de nuestra propuesta respecto de aquellas.

1.5.1. Utilización del Enfoque MDD

Los trabajos relacionados al diseño de estructuras de datos que utilizan el enfoque **MDD** presentados por otros autores, plantean los siguientes objetivos: mejorar la productividad en el desarrollo de un **DW**, en el marco **MDA**, ([MT09], [MOT07], [MTL06], [ZC06], [MTSP05]); utilizar el enfoque **MDA** en el diseño de **DW** espaciales ([GT08]); considerar aspectos de seguridad en el **DW** ([STFP07a], [STFP07b], [STFP07c], [Sol+07], [STBF09]) e implementarlos en una herramienta **OLAP** específica ([Bla+09]) o, por último, utilizar el enfoque **MDA** para el desarrollo de un **ORDB** ([VVCM07]).

El enfoque utilizado en todos los casos por los trabajos presentados es en el marco de **MDA**, esto implica el uso de estándares asociados propuestos por la **OMG**, (**UML** y *profiles*, **OCL**, **XMI**, **CWM**, **QVT**).

En nuestra propuesta, el enfoque utilizado es **MDD**, en particular **DSM**, creamos modelos específicos del dominio utilizando un lenguaje focalizado y específico para cada uno de ellos (**DSL**); en particular, no utilizamos **CWM** sino metamodelos más simples, instancias de **MOF**, para los modelos de datos, el modelo **MD** y el **RM**; además, diseñamos metamodelos específicos para los modelos utilizados en el proceso: para la construcción del **AG**, el metamodelo **AG** y para la construcción del **QG**, el metamodelo **QG**. No usamos **UML** ni *profiles* para el diseño de **PIM** del modelo de datos fuente, ya que consideramos que el modelo **ER** es más expresivo para el modelado de datos. Aunque sí utilizamos **OCL** para establecer restricciones sobre los metamodelos propuestos.

Respecto de las estructuras de almacenamiento, se diferencia de los trabajos referenciados, principalmente, en el modelo **MD** propuesto: el **HDW** representa una nueva estructura de datos

que combina e integra, en un solo modelo, un **DW** y un **HDB**; este modelo incluye, además del hecho principal de análisis, estructuras temporales vinculadas a los niveles de las jerarquías dimensionales que posibiliten registrar los datos y recuperar la información que varíase en el tiempo.

1.5.2. Consultas Gráficas Automatizadas

Respecto a las consultas gráficas, los trabajos relacionados pueden clasificarse en: aquellos vinculados con consultas a base de datos **MD** ([RTTZ08], [MS06], [TSH01]); a los lenguajes relacionadas con consultas en **TDB** ([FKSS06]); a los lenguajes vinculados a consultas en **BD** ([Rei02], [ON01], [KG95]); por último, aquellos vinculados al diseño de interfaces visuales utilizados para el aprendizaje, tanto en la creación de consultas ([ACDS02]), como para la comprensión del proceso subyacente del **DBMS** ([AYW08])

Dentro de esta clasificación, nuestra propuesta es, en parte, una conjunción de las dos primeras: la interface gráfica presentada le permite al usuario final realizar consultas **MD** e históricas en forma automática.

No obstante, lo más significativo que diferencia nuestra propuesta del resto es el diseño de un entorno gráfico, derivado automáticamente del **HDW**, que permite la generación automática de sentencias **SQL** para realizar consultas **MD** e históricas. Hasta donde hemos visto, no hemos encontrado investigaciones que utilicen el enfoque **MDD**, ni para la derivación de un entorno gráfico de consultas, ni así tampoco para la creación de sentencias **SQL**, en forma automática, sobre una estructura **TMD**.

1.6. Trabajos Futuros

A partir de la tesis presentada, se abre un abanico de posibles líneas de investigación asociadas que no fueron consideradas en el desarrollo de la tesis pero que ameritan ser tenidas en cuenta en futuros trabajos.

Detallaremos, a continuación, los temas que no hemos considerados y cuya solución implica una línea de investigación a desarrollar:

- **Creación de un DLS para transformaciones de DW:** en este trabajo hemos utilizado el lenguaje **ATL** para definir las transformaciones entre los modelos propuestos. **ATL** es un lenguaje específico del dominio de las transformaciones, es decir, provee construcciones sintácticas focalizadas en la definición de transformaciones entre modelos. Adicionalmente estos lenguajes de transformación (LT) pueden admitir un grado más de especialización, es decir, podríamos definir lenguajes de transformación específicos de dominio (LTED). En nuestro caso, podríamos definir un LT específico para transformaciones de **DW**. Contar con un lenguaje específico versus un lenguaje más general, como **ATL**, facilitaría notablemente la definición y la reutilización de las transformaciones.
- **Verificación formal de las propiedades de las transformaciones:** Se espera que cada transformación realice solamente modificaciones sintácticas sobre los modelos, pero respetando su semántica. La verificación de esta propiedad (semantic preserving transformations) de las transformaciones es un tema muy complejo; sin embargo, al restringirlo al dominio específico de los **DW** sería posible obtener mejores resultados.
- **Creación de una interface de consulta eficiente:** la interface gráfica presentada en el método de diseño y posteriormente implementada en el prototipo desarrollado en

ECLIPSE es, sin duda, *naif*. El uso de interfaces gráficas implica consideraciones sobre tipo de iconos a utilizar, diferentes colores, distribución de los iconos en la pantalla, etc. que no han sido considerados en el actual trabajo y que deberán considerarse en ulteriores investigaciones.

- **Considerar requisitos del usuario en el diseño del HDW:** nuestro trabajo utiliza como modelo fuente, para el proceso de transformaciones, un modelo de datos conceptual, expresado en un **ER**, del cual consideramos que representa los requisitos de información de los usuarios, al menos en lo referente a la aplicación transaccional. No hemos considerado cómo evaluar y plasmar en el diseño del **HDW** los requisitos de información del usuario en aspectos **MD** y temporales. En parte esto es así ya que las transformaciones comienzan por el **PIM**, sin considerar el **CIM**, en forma explícita. La transformación de **CIM** a **PIM**, es un ámbito no muy desarrollado, que abre una línea de investigación a considerar en futuros trabajos.
- **Ampliación del modelo de datos:** el modelo de datos temporal presentado utilizado es el modelo **ER** estándar, donde solo se incluyen las construcciones básicas y, mediante ellas, la captura, en forma implícita, de los aspectos temporales. No hemos considerado inicialmente en nuestro modelo de datos interrelaciones de grado > 2 y tampoco extendimos su semántica con conceptos tales como generalizaciones, agregaciones y construcciones temporales. Estos aspectos ameritan ser evaluados y considerados en futuras ampliaciones del modelo de datos utilizado.
- **Restricciones de Integridad Temporal:** el modelo, tal cual está planteado, no contempla restricciones con respecto a las actualizaciones, pudiendo darse el caso de solapamientos temporales. El establecimiento de restricciones con respecto a la inserción, borrado y modificación de valores temporales debería impedir posibles inconsistencias en la base de datos. Además, como se plantea en el capítulo 3, es necesario considerar el tiempo válido de las entidades temporales como subconjunto de los tiempos válidos de atributos, entidades e interrelaciones involucradas. Estos temas constituyen un ámbito de investigación a considerar.
- **Derivación automática del proceso ETL:** otro aspecto que no hemos considerado en nuestro diseño del **HDW** es cómo realizar el proceso de **ETL**. Este proceso es importante porque es el encargado de integrar datos de diferentes fuentes heterogéneas. Para la construcción del modelo conceptual partimos de un modelo **ER** que representa, a una base de datos no histórica. La ampliación a un modelo temporal no implica, conceptualmente, mayores inconvenientes. La carga de datos históricos sí requerirá considerar estrategias que permitan el poblado del **HDW** a partir de datos provenientes de copias de resguardo almacenados en distintos soportes y formatos. Una línea de investigación a considerar es, en el contexto de **MDD**, la generación automática de código de procesos **ETL**.
- **Integración de la herramienta de transformación con DBMS's:** el prototipo presentado culmina su proceso de transformación con la generación, por un lado, de sentencias **SQL** para la creación de estructuras de almacenamiento y, por otro lado, en sentencias **SQL** que resuelven las consultas sobre el **HDW**; ambos tipos de sentencias no son ejecutadas directamente sino que, el texto generado, debe ser ejecutado posteriormente en un **DBMS**. Por lo tanto, otra línea de investigación a seguir es la

integración directa de la herramienta de transformación desarrollada en ECLIPSE con un **DBMS**.

- **Uso de un PSM Objeto Relacional:** hemos utilizado el **RM**, en particular el estándar **SQL92** para el desarrollo del **PSM**. Las bases de datos **ORDB** (el estándar **SQL2003**) presentan construcciones tales como los tipos abstractos de datos definidos por el usuario, que permitirían una más simple representación el modelo temporal propuesto, permitiendo desarrollar una línea de investigación asociada.

1.7. Publicaciones Vinculadas a la Tesis

A continuación, se listan los principales trabajos publicados cuya temática está vinculada al de los objetivos planteados en la tesis:

- Carlos Neil, Jerónimo Irazábal, Marcelo De Vincenzi, Claudia Pons. **Graphical Query Mechanism for Historical DW within MDD**. XXIX Conferencia Internacional de la Sociedad Chilena de Ciencia de la Computación (IEEE Press). Chile 2010
- Neil, Carlos, Pons Claudia. **Aplicando QVT en la Transformación de un Modelo de Datos Temporal**. Jornadas Chilenas de Computación. Punta Arenas. Chile. 2008
- Neil Carlos, Baez Martín, Pons Claudia. **Usando ATL en la Transformación de Modelos Multidimensionales Temporales**. XIII Congreso Argentino de Ciencias de la Computación. Corrientes y Resistencia, Argentina. 2007.
- Neil Carlos, Pons Claudia. **Aplicando MDA al Diseño de un Data Warehouse Temporal**. VII Jornada Iberoamericana de Ingeniería de Software e Ingeniería del Conocimiento. Lima, Perú. 2007.
- Neil Carlos, Pons Claudia. **Diseño Conceptual de un Data Warehouse Temporal en el Contexto de MDA**. XII Congreso Argentino de Ciencias de la Computación. CACIC. San Luis. Argentina. 2006.
- Neil Carlos, Ale Juan. **A Conceptual Design for Temporal Data Warehouse**. 31° JAIIO. Santa Fe. Simposio Argentino de Ingeniería de Software. 2002.

1.8. Referencias

- [ACDS02] Aversano, L., G. Canfora, A. De Lucia, S. Stefanucci. "Understanding SQL through Iconic Interfaces," *Proceedings 26th Annual International Computer Software and Applications Conference*, 2002, pp. 703-708.
- [AYW08] Brett Allenstein , Andrew Yost , Paul Wagner , Joline Morrison, A query simulation system to illustrate database query execution, *Proceedings of the 39th SIGCSE technical symposium on Computer science education*, March 12-15, 2008, Portland, OR, USA
- [Bla+09] Blanco, C., Garca-Rodriguez de Guzmán, I., Fernández-Medina, E., Trujillo, J., Piattini, M.: "Applying QVT in order to implement Secure Data Warehouses in SQL Server Analysis Services", *Journal of Research and Practice in Information Technology*, 41, 2 (2009), 119-138.

- [CD97] Chaudhuri, Surajit and Dayal, Umesh An Overview of Data Warehousing and OLAP Technology, ACM SIGMOD Record 26(1), March 1997.
- [EC00] Eder, Johann. Concilia, Christian. Evolution of Dimension Data in Temporal Data Warehouses Technical Report, 2000.
- [EKM01] Eder, J. Koncilia, C. and Morzy, T.. A Model for a Temporal Data Warehouse. In Proc. of the Int. OESSEO 2001 Conference, Rome, Italy, 2001.
- [FKSS06] Fails, J., Karlson, A., Shahamat, L., Shneiderman, B. A Visual Interface for Multivariate Temporal Data: Finding Patterns of Events over Time Proceedings of IEEE Symposium on Visual Analytics Science and Technology (VAST 2006), 167-174. 2006.
- [GHRU97] Gupta, H. Harinarayan, V. Rajaraman, A. and Ullman, J. Index Selection for OLAP. Proceeding ICDE '97. 1997.
- [GMR98a] Golfarelli, M. Maio, D. Rizzi, S. Conceptual Design of Data Warehouses from E/R schemes, Proceedings 31st Hawaii International Conference on System Sciences, 1998.
- [GT08] Glorio O and Trujillo J., "An MDA Approach for the Development of Spatial Data Warehouses," in DaWaK, Turin, Italy, 2008, pp. 23–32.
- [HMV99] Hurtado, C., Mendelzon, A. & Vaisman, A., Maintaining Data Cubes Under Dimension Updates, in 'Proc. of the 15th Int. Conf. on Data Engineering', pp. 346-355. 1999.
- [HRU96] Harinarayan, V., Rajaraman, A. & Ullman, J.D. Implementing data cubes efficiently. ACM SIGMOD Record, 25(2): 205--216. 1996.
- [Inm02] Inmon, W. Building the Data Warehouse. John Wiley & Sons, 2002.
- [Kim96] Kimball, R. The Data Warehouse Toolkit. John Willey and Sons, 1996.
- [KG95] Kouramajian, V. and Gertz, M., A graphical query language for temporal databases. in Proceedings of 14th International Conference on Object-Oriented and Entity Relationship Modeling, (1995), Springer-Verlag, 388-399.
- [KWB03] Kleppe, Anneke G. and Warmer Jos, and Bast, Wim. MDA Explained: The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
- [LJ09] B. Liu and H. Jagadish. A spreadsheet algebra for a direct data manipulation query interface. In ICDE, 2009.
- [MOT07] Mazón, J.N., Ortega, E., Trujillo, J.: Ingeniería inversa dirigida por modelos para el diseño de almacenes de datos. In: JISBD. 2007.

- [MS06] Svetlana Mansmann and Marc H. Scholl. Extending visual OLAP for handling irregular dimensional hierarchies. In *DaWaK*, pages 95 (105, 2006).
- [MT09] Jose-Norberto Mazon and Juan Trujillo. A Hybrid Model Driven Development Framework for the Multidimensional Modeling of Data Warehouses. Jose-Norberto Mazon and Juan Trujillo. SIGMOD Record, June 2009 (Vol. 38, No. 2).
- [MTL06] Jose-Norberto Mazón, Juan Trujillo, Jens Lechtenböcker: A Set of QVT Relations to Assure the Correctness of Data Warehouses by Using Multidimensional Normal Forms. ER 2006: 385-398.
- [MTSP05] Mazón Jose Norberto, Trujillo Juan, Serrano Manuel, Piattini Mario: Applying MDA to the Development of Data Warehouses. DOLAP 2005: 57-66.
- [NA02] Neil Carlos, Ale Juan. A Conceptual Design for Temporal Data Warehouse. 31° JAIIO. Santa Fe. Simposio Argentino de Ingeniería de Software. 2002.
- [ON01] Owei, V. and Navathe, S. B. (2001) Enriching the conceptual basis for query formulation through relationship semantics in databases. *Inf. Syst.*, 26, 445–475.
- [PGP09] Claudia Pons, Roxana Giandini, Gabriela Pérez. Desarrollo de Software Dirigido por Modelos. Conceptos teóricos y su aplicación práctica. cantidad de pgs. 300. Editorial: McGraw-Hill Education. 2009.
- [RA08] Oscar Romero, Alberto Abelló: MDBE: Automatic Multidimensional Modeling. ER 2008: 534-535.
- [Rei02] Reiss SP (2002) A visual query language for software visualization. In: Proceedings of the IEEE 2002 Symposia on Human Centric Computing Languages and Environments (HCC'02), pp 80–82.
- [RTT08] Ravat, Franck; Teste, Olivier; Tournier, Ronan; Zurfluh, Pilles. Algebraic and Graphic Languages for OLAP Manipulations International Journal of Data Warehousing and Mining, Vol. 4, Issue 1. 2008.
- [Sol+07] Soler Emilio, Trujillo Juan, Fernández-Medina Eduardo, Piattini Mario: A set of QVT relations to transform PIM to PSM in the Design of Secure Data Warehouses. ARES 2007: 644-654.
- [STBF09] Emilio Soler, Juan Trujillo, Carlos Blanco, Eduardo Fernández-Medina. Designing Secure Data Warehouses by Using MDA and QVT Journal of Universal Computer Science, vol. 15, no. 8 (2009), 1607-1641.
- [STFP07a] Soler, E., Trujillo, J., Fernández-Medina, E., Piattini, M.: Una extensión del metamodelo relacional de CWM para representar Almacenes de Datos Seguros a nivel lógico. In: JISBD. (2007).
- [STFP07b] Soler E., Trujillo J., Fernández-Medina E., Piattini M.: Aplicación de QVT al Desarrollo de Almacenes de Datos Seguros: Un Caso de Estudio. IDEAS 2007. Isla Margarita (Venezuela).

- [STFP07c] Soler, E., Trujillo, J., Fernández-Medina, E., Piattini, M.: Un Conjunto de Transformaciones QVT para el Modelado De Almacenes de Datos Seguros. In: JISBD Workshops (DSDM), (2007).
- [TBJS99] Tryfona, N. Busborg, F. Christiansen, J.G.B. starER: A Conceptual Model for Data Warehouse Design. In ACM Second International Workshop on Data Warehousing and OLAP (DOLAP'99), pp.3-8, November 1999, Missouri, USA.
- [TSH01] D. Tang C. Stolte and P. Hanrahan, "Polaris: A system for query, analysis and visualization of multi-dimensional relational databases," Transactions on Visualization and Computer Graphics, 2001.
- [VVCM07] Vara, J. M. Vela, B. Cavero, J. M. Marcos, E. Transformación de Modelos para el Desarrollo de Base de Datos Objeto-Relacionales. IEE latin American transactions vol 5. No4, july 2007.
- [ZC06] Zepeda, L., Celma, M. Aplicando MDA al Diseño Conceptual de Almacenes de Datos. 9º Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes de Software (IDEAS'06). Mar del plata, Argentina, 2006.