

Construcción de un sistema de apoyo a la toma de decisiones para el área gerencial del Hospital de Clínicas

Alejandro Gutiérrez[‡], Regina Motz, Beatriz Revello, Lydia Silva

Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

Resumen

Se presenta la aplicación de una metodología para la construcción de una base de datos para satisfacer requerimientos de la toma de decisiones (data warehouse) para el área gerencial del Hospital de Clínicas del Uruguay. El Hospital de Clínicas se presenta como un buen exponente de la variedad de estados en que pueden encontrarse los sistemas legados existentes en una institución de su porte, y las dificultades que se presentan a la hora de reconciliarlos para obtener un data warehouse corporativo. En este trabajo se analizan las etapas necesarias para la construcción del data warehouse, se propone el uso de un modelo conceptual multidimensional para definir las estructuras de datos que respondan a los requerimientos del área gerencial del hospital y se presenta una solución para la carga y refresco del mismo. El desarrollo de este tipo de sistema de información es relativamente nuevo en el área médica. Este trabajo aporta una identificación de etapas a considerar y una cuantificación de sus tiempos obtenidos en la experiencia de resolver requerimientos del área gerencial de dicho Hospital.

Palabras claves:

Sistemas de información; Diseño conceptual y lógico; Modelos multidimensionales; Bases de datos para la toma de decisiones; Sistemas legados; Gestión administrativa hospitalaria.

1. Introducción

Durante años las instituciones y empresas han acumulado la información que los sistemas generan al resolver los procesos (transacciones) que hacen posible la gestión cotidiana. Esta información es sumamente valiosa a la hora de tomar decisiones gerenciales, que necesitan conocimientos sobre datos históricos y tendencias.

El caso típico de una persona buscando información estratégica consiste en obtener cierta información, y luego utilizar el resultado obtenido para dirigir su próxima

consulta, por lo que se debe contar con una herramienta que permita esa interacción en un corto tiempo y minimizando la interferencia con los sistemas que realizan la gestión cotidiana.

El uso de herramientas tradicionales basadas en el desarrollo de aplicaciones sobre sistemas de gestión de bases de datos para obtener información estratégica para la toma de decisiones presenta inconvenientes de diferente índole, como tiempos de respuesta extremadamente lentos, dificultad o imposibilidad de expresar una consulta compleja, y dificultad para la obtención de resultados que recopilen información de fuentes diversas.

Las planillas electrónicas si bien son muy amigables para el usuario final, presentan problemas de eficiencia y no son confiables para grandes volúmenes de información.

Una solución encontrada a estos inconvenientes en los últimos años ha sido la construcción de sistemas de data warehouses [1] que almacenan la información en estructuras de datos especiales, apropiadas para favorecer la eficiencia en las operaciones de consulta, y que con herramientas adecuadas de análisis de datos (por ejemplo OLAP *On-Line Analytical Processing*), ofrecen facilidades de apoyo en la toma de decisiones.

Un data warehouse puede definirse como una base de datos que almacena información para satisfacer requerimientos de toma de decisiones. Una característica diferenciadora de este tipo de base de datos es que su información proviene de sistemas operacionales ya existentes conocidos como sistemas legados. y debe ser mantenida a medida que la información de los sistemas legados es agregada y/o modificada.

El Hospital de Clínicas Dr. Manuel Quintela del Uruguay se presenta como un buen exponente de la variedad de estados en que pueden encontrarse los sistemas legados existentes en una institución de su porte, y las dificultades que se presentan a la hora de reconciliarlos para obtener un data warehouse corporativo. También presenta la riqueza de una institución que abarca diversos servicios en el área de la salud, con la variedad de enfoques de análisis que eso propicia.

[‡] El autor ha recibido apoyo de la Comisión Sectorial de Investigación Científica de la Universidad de la República, Uruguay



Figura 1. Áreas administrativas del Hospital de Clínicas

Dicho hospital es una institución pública abocada a la atención sanitaria de la población, y a la docencia universitaria en el área de la salud y en todo nivel de especialización. Los procedimientos y sistemas informáticos del Hospital son un ejemplo de la variedad de estados en que pueden encontrarse los sistemas existentes en una institución de más de 50 años de historia con capacidad de hasta 800 camas. Su gestión se divide en dos grandes áreas administrativas que se ilustran en la Figura 1.

Los *Servicios Asistenciales* comprenden la operativa clínica y los servicios que se dan al público, tanto generales (consulta ambulatoria, hospitalización) como especializados (servicios médicos proporcionados a otras instituciones por contrato). Se agrupan en 3 servicios: Clínicos, Diagnóstico y Tratamiento y Técnicos no médicos. La principal tarea en la administración de esta área es controlar la actuación de los profesionales y técnicos, así como resolver problemas médico administrativos y de docencia.

Los *Servicios de Apoyo* comprenden el manejo administrativo del hospital, incluyendo el control de compras y producción interna de artículos de uso de la institución, administración de recursos humanos y materiales. Estos servicios sustentan al hospital en desempeño de la misma para brindar los Servicios Asistenciales. Se agrupan en 2 servicios: Generales y Administración General. La principal tarea en la administración de esta área es planificar gastos y pagos hasta el límite de la asignación presupuestal y realizar el control necesario sobre ellos.

En este artículo se presenta la aplicación de una metodología para la construcción de un data warehouse para el área de Servicios de Apoyo del Hospital de Clínicas.

El resto del artículo se organiza de la siguiente manera. En la sección 2 se describe la metodología de diseño usada en el desarrollo del sistema de data warehouse. En la sección 3 se muestra el resultado de la aplicación de dicha metodología al desarrollo de una parte del mismo para el Hospital de Clínicas. Por último en la Sección 4 se presentan las conclusiones incluyendo el trabajo futuro.

2. Metodología de diseño

La metodología de diseño utilizada [2] se basa en la

definición de un data warehouse como la combinación de estructuras de datos especializadas para resolver requerimientos de análisis de información en distintas áreas de una institución. La estructura especializada se puede presentar como una función que explica una medida en términos de parámetros llamados dimensiones.

Para la representación de esta función usamos un modelo multidimensional que consiste en la visualización de la información en una estructura de matriz multidimensional denominada Cubo. Cada cubo contiene la información correspondiente a un objeto de la realidad que interesa analizar. Los ejes representando a las dimensiones contienen los valores que, seleccionados por el usuario, condicionarán el análisis de la información. Las celdas contienen los datos a medir llamados hechos o medidas, que podrán ser por ejemplo sumados y promediados. Los análisis se obtienen en forma de cifras, listados o gráficos.

Las estructuras multidimensionales de las distintas áreas combinadas darán origen al esquema de un data warehouse. Esta combinación es básicamente una unión en la cual se comparten las dimensiones comunes de las diferentes estructuras. La información de los sistemas legados se recolecta, depura e integra para generar la información a almacenar en el data warehouse.

Por lo tanto, desarrollar un data warehouse implica la definición del tipo de información que se desea obtener del mismo, cuales son los sistemas legados que contienen dicha información, que estructura tendrán los depósitos de datos a usar, así como las operaciones a aplicar sobre los datos de dichos sistemas legados para poblar adecuadamente la estructura del data warehouse.

A continuación se detalla el flujo de las etapas de trabajo y los documentos producidos en cada una. La Figura 2 es una representación gráfica de los documentos producidos en las etapas de la metodología y de la relación entre ellos.

Etapa 1. Entrevistas

Las actividades de esta etapa incluyen la definición de requerimientos a satisfacer por el sistema a construir y el relevamiento de la realidad existente en los depósitos de datos de los sistemas legados que dispone la institución. El material consultado en la etapa fue [3, 4].

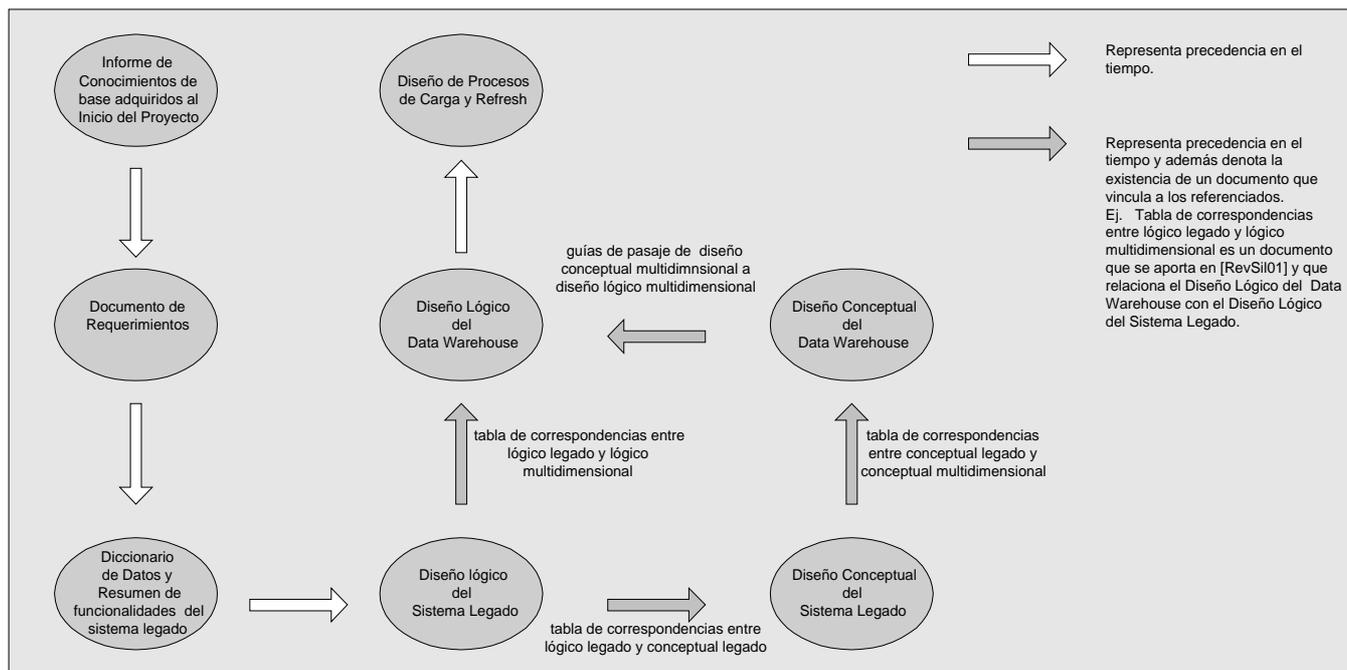


Figura 2. Relación de documentos implicados en la Metodología de Diseño

Las actividades y documentos relacionados son:

- Obtención de requerimientos a partir de entrevistas con el nivel gerencial de la institución. El resultado de esta actividad es el documento de requerimientos.
- Relevamiento de los sistemas legados, revisión de la documentación existente y *entrevistas con personal de dichos sistemas*. El resultado de esta actividad son los siguientes documentos:
 - diccionario de datos
 - diseño lógico del sistema legado
 - resumen de funcionalidades del sistema legado

Etapa 2. Análisis

La segunda etapa consiste en la etapa de análisis en la cual se analizan los documentos obtenidos y/o generados en la etapa anterior, y a partir de la síntesis conceptual de los mismos y nuevas *entrevistas con personal de los sistemas legados*, se construyen el *diseño conceptual del sistema legado* y la *tabla de correspondencias entre el diseño lógico y conceptual del sistema legado*, para cada uno de los sistemas legados involucrados. Estos documentos son útiles para la comprensión del carácter semántico de la información almacenada en los sistemas legados.

Etapa 3. Diseño Conceptual Multidimensional

La tercer etapa se enfoca en la construcción del diseño conceptual del data warehouse a partir de los

requerimientos identificados en la Etapa 1 y la documentación existente del sistema legado. El principal material consultado para la definición de esta etapa fue [5, 6, 7, 8]. Las actividades y documentos relacionados con esta etapa son:

- Elección de la herramienta a usar para representar el diseño conceptual del modelo de datos del data warehouse
- Construcción del *diseño conceptual multidimensional*.
- Construcción de la *tabla de correspondencias entre el diseño conceptual del sistema legado y el diseño conceptual del data warehouse*

Etapa 4. Diseño Lógico Multidimensional

La cuarta etapa abarca la construcción del diseño lógico multidimensional que se basa fundamentalmente en el diseño conceptual multidimensional construido en la etapa anterior y los requerimientos multidimensionales identificados. El principal material consultado para la definición de esta etapa fue: [3, 7, 8, 9, 10]. Las actividades y documentos relacionados con la etapa son:

- Definición de la arquitectura de los depósitos de datos del data warehouse.
- Definición del tipo de almacenamiento dependiendo del uso de tecnología relacional o multidimensional. En el caso del uso de tecnología relacional esta actividad incluye el diseño del esquema relacional del depósito de datos del data warehouse.

- Diseño de los procesos de carga y refresco que poblarán al data warehouse, que implican estrategias de extracción, limpieza, notificación de errores y administración de cambios detectados en los datos provenientes de los sistemas legados.

El proceso de carga inicial del data warehouse genera las instancias de las estructuras de datos del mismo con la información proveniente de los sistemas legados, depurada, integrada y eventualmente resumida, existente en el momento de ejecución del mismo

El proceso de refresco tiene como cometido la actualización de la información del data warehouse con los nuevos datos generados o modificados en los sistemas legados a través del tiempo con la frecuencia adecuada. El refresco de un data warehouse determina el uso efectivo de los datos recolectados y resumidos desde los orígenes. La calidad de los datos provistos a quienes toman decisiones, depende de la capacidad del data warehouse de propagar los cambios hechos en los sistemas legados en el tiempo definido como conveniente para que la aplicación tenga datos considerados “frescos” en todo momento.

Los documentos construidos en esta etapa son los siguientes:

- Diseño lógico multidimensional.
- Tabla de correspondencias entre el diseño lógico del sistema legado y el diseño lógico del data warehouse.
- Diseño de procesos de carga y refresco.

Etapa 5. Implementación

En esta última etapa se selecciona la herramienta a usar en la implementación de los procesos de carga y refresco del data warehouse. y se realiza además la implementación de dichos procesos.

Como resultado de la ejecución de los procesos de carga y refresco se obtienen las estructuras de datos del data warehouse convenientemente actualizadas y se genera un *resumen de errores* detectados durante el proceso de carga o refresco. El uso de este resumen de errores facilita la corrección de los datos en los sistemas legados.

3. Aplicación de la metodología

La metodología presentada fue utilizada para resolver los requerimientos del área de Servicios de Apoyo del Hospital de Clínicas. En esta sección presentamos un ejemplo simplificado extraído del trabajo [2] para ilustrar la forma en que se aplicó la metodología. Luego, mostramos los tiempos dedicados a cada etapa. Finalmente, realizamos una breve discusión sobre algunas causas y dificultades encontradas que explican los tiempos dedicados y que a su vez permiten sugerir guías para futuros trabajos.

3.1 Ejemplo

Presentaremos la aplicación de la metodología limitándonos a los requerimientos asociados a la evolución de compra de un artículo en un período de tiempo dado. El resultado final será la construcción de un cubo denominado Compras

Etapa 1. Entrevistas

Los requerimientos a satisfacer se expresan como (1) un conjunto de consultas concretas, que en conjunto identifican el área de análisis que deberá cubrir el data warehouse y (2) un conjunto de cuantificadores de éxito de la gestión administrativa que se abarca.

Un ejemplo de consulta se presenta a continuación.

Para un proveedor, mostrar la evolución del porcentaje de órdenes de compra que ha cumplido respetando los plazos de entrega en un período de tiempo dado.

Un ejemplo de cuantificador de éxito es el *índice de calidad* de un artículo que se define como:

el cociente entre el valor de las devoluciones por baja calidad y el valor total del artículo recibido.

En el relevamiento de los sistemas legados se evidencia la inexistencia de documentación en todos, y la carencia de información almacenada electrónicamente en algunos, por lo que mediante un gran número de *entrevistas con el personal de los sistemas legados* se redefine el alcance de los sistemas a abarcar en el trabajo y se genera el diccionario de datos, diseño lógico, y resumen de funcionalidades para el único sistema factible de ser incluido.

Etapa 2. Análisis

Se adquirió un conocimiento más detallado de la realidad a modelar mediante el análisis de la documentación construida en la etapa anterior y nuevas *entrevistas con el personal del sistema legado*.

Se construyó el Diseño Conceptual del Sistema de Recursos Materiales y control del Gasto y la tabla de correspondencias entre diseños lógico y conceptual de dicho sistema legado.

Etapa 3. Diseño Conceptual Multidimensional

Para el diseño conceptual multidimensional se utilizó CMDM (*Conceptual Multidimensional Data Model*) definido en [5]. CMDM es un modelo conceptual orientado al diseño de estructuras multidimensionales que propone estructuras de datos y un mecanismo de especificación de restricciones de integridad.

Una característica interesante de CMDM es que distingue entre *dimensiones* que identifican objetos de la realidad y *relaciones dimensionales* que representan las relaciones multidimensionales existentes entre dichos objetos.

Como ejemplos de dimensiones, citamos a Tiempo, Ofertas, Autorizaciones para Gastar (APG), Artículos, Proveedores y Transacciones de Ofertas. Como ejemplos de relaciones dimensionales, citamos *RelOferta* y *Ordenes*, las cuales se ilustran en las Figuras 3 y 4, respectivamente.

La relación dimensional *RelOfertas* vincula todos los elementos participantes de la Oferta que un proveedor hace por la venta de un artículo en el marco de una Autorización para Gastar que emite la institución.

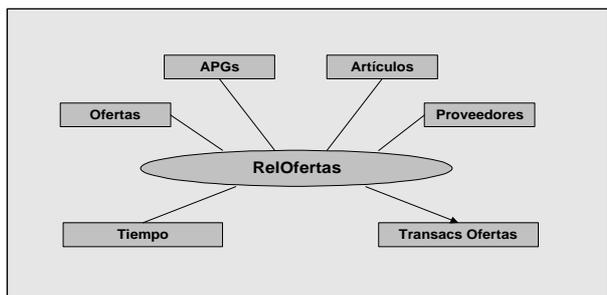


Figura 3. Relación Dimensional RelOfertas

La relación dimensional *Ordenes* vincula todos los elementos participantes de la Orden de Compra que la institución emite en respuesta a una Oferta que ha sido aceptada.

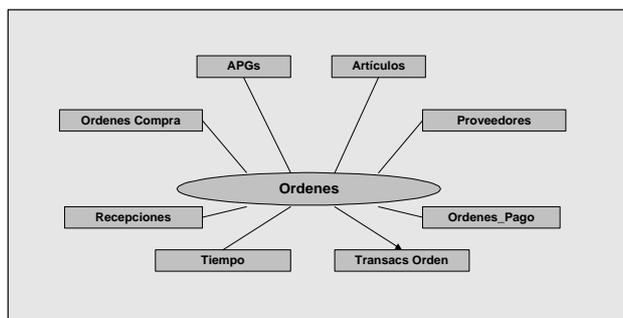


Figura 4. Relación Dimensional Ordenes

En las Figuras 3 y 4 se destaca con una flecha orientada hacia una de las dimensiones que la dimensión apuntada representa la medida de la estructura.

Se construyó además la *tabla de correspondencias entre el diseño conceptual del Sistema De Recursos Materiales y Control del Gasto y el diseño conceptual del data warehouse*

Etapa 4. Diseño Lógico Multidimensional

En esta etapa se definió un data warehouse global con el nivel más bajo de granularidad en los datos. Se considera que los datos están al nivel más bajo de granularidad

cuando no contienen resúmenes u otras operaciones de agrupación realizadas sobre los datos provenientes de los sistemas legados.

Se realizó un cálculo de tamaño relativo al momento de la finalización del proceso de carga y se proyectó el crecimiento hasta 5 años. A modo de ejemplo presentamos el volumen manejado por el Hospital.

- Tamaño al momento de la carga inicial: 300 MB
- Tamaño proyectado en 5 años: 1,5 GB

Se consideró el uso de tecnología relacional o multidimensional. Según consideraciones de tamaño y grado de dispersión de los datos existentes, se concluyó conveniente adoptar: tecnología relacional para el almacenamiento de más bajo nivel de granularidad y tecnología multidimensional para los resúmenes precalculadas. En particular para el almacenamiento de los datos de más bajo nivel de granularidad mediante tecnología relacional se diseñó una tabla de hechos ilustrado en Figura 5 en el que confluyen las dos relaciones multidimensionales ilustradas en Figuras 3 y 4.

Dicha confluencia fue necesaria por la alta frecuencia prevista de operaciones entre las dos relaciones multidimensionales (conocidas como *drill across*) que serían necesarias para evacuar los requerimientos planteados en la Etapa 1.

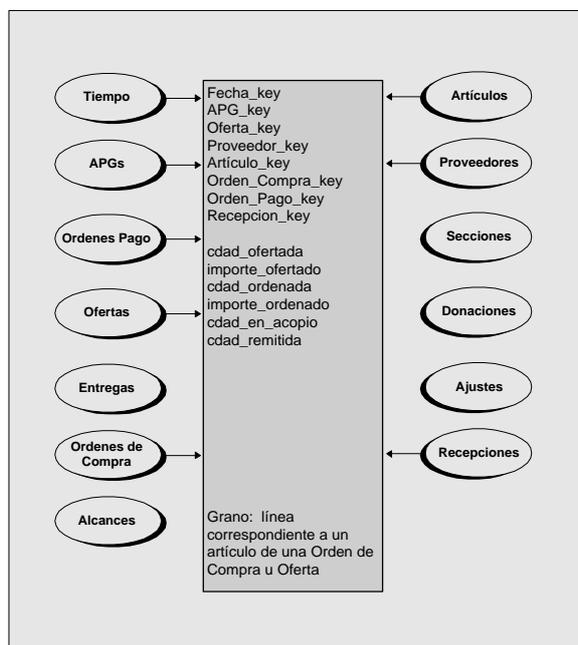


Figura 5. Tabla de hechos Compras

El proceso de carga de dimensiones, y tablas de hechos comprende cuatro etapas. A saber,

- Extracción de datos del sistema legado.
- Almacenamiento de los datos en el área de trabajo

del data warehouse

- Depuración de los datos.
- Carga de las estructuras finales de las dimensiones y tablas de hechos.

La frecuencia de ejecución del proceso de refresque es diaria para optimizar el grado de frescura de los datos a usar. Dado que el tiempo completo de refresque es un lapso pequeño de tiempo de aproximadamente 20 minutos resulta perfectamente viable su realización diaria.

El método de extracción de información de los sistemas legados hace uso de snapshots (volcado completo en archivos de texto). Se consideró esta opción dado que el sistema legado, por desempeño sólo mantiene datos de dos años, y que el volumen de estos datos resulta razonablemente pequeño para ser bajado en una corta ventana de tiempo, se encontró entonces justificable preferir bajar como archivo de texto todo el contenido de tablas, y realizar consultas SQL que identificarán las tuplas a considerar para refrescar el Data Warehouse. La detección de los cambios se realiza entonces por comparación de snapshots.

La estrategia de actualización del data warehouse con los cambios detectados es diferente para las tablas de hechos y para las dimensiones.

En las tablas de hechos se sobrescribe la tupla modificada. Se siguió esta opción porque se observó que las actualizaciones que repercuten en cambios en tablas de hechos tienen su origen en la corrección de errores ocurridos en el sistema legado. Por consiguiente, se sustituyen las tuplas correspondientes para reflejar la situación existente en dicho sistema.

En las dimensiones a su vez se siguen estrategias diferentes que se describirán a continuación.

Para las dimensiones APG, Artículos, Ordenes de Compra y Proveedores, la estrategia consiste en agregar una nueva tupla a la dimensión para el mismo ítem, con la nueva versión de los atributos, de manera tal que al asociarse nuevos hechos, se vinculen a la versión más reciente de la tupla en la dimensión. Los hechos anteriores a la actualización, permanecerán vinculados a la tupla más antigua de la dimensión. Esto responde al interés manifestado por los usuarios de conservar los valores existentes previo a la actualización de las referidas dimensiones, y también actualizar al data warehouse con los nuevos.

Para las restantes dimensiones, la estrategia consiste en sobrescribir la tupla de la dimensión con los nuevos datos logrando así que tanto los hechos previos como posteriores al cambio resulten vinculados a la nueva descripción de dimensión. Esto se corresponde con el interés del usuario de no recordar los valores anteriores en esas dimensiones, ya sea porque se asume que los cambios se debieron a errores de digitación, o porque carezca de relevancia recordar si antes tuvieron otro valor.

Etapas 5. Implementación

Para la implementación de los procesos de carga y refresque se utilizó una herramienta ETL (Extraction, Transformation and Loading.)

Esta categoría de herramientas está orientada a la programación de dichos procesos. Consideramos el uso de estas herramientas en lugar de realizar una codificación específica debido a dos facilidades principales que ellas ofrecen. Por un lado, la generación automática de metadatos (es decir, datos acerca de la ejecución de los procesos de carga y refresque de los datos) y por otro lado las facilidades para el mantenimiento de los programas de carga y refresque.

En este trabajo se experimentó con DTS (*Data Transformation Services*) de Microsoft Corporation [11].

3.2 Tiempo de las etapas

La Tabla 1 resume los tiempos dedicados a cada una de las etapas en el desarrollo del data warehouse teniendo en cuenta el servicio de Compras. Las horas indicadas corresponden a horas efectivas realizadas por 2 personas.

Se realizan algunas consideraciones sobre las Etapas 1, 2 y 5. En la Etapa 1 se incluye la preparación previa y documentación posterior de las entrevistas así como el tiempo de análisis entre entrevistas.. En la Etapa 2 se incluye entrevistas aclaratorias y de confirmación del modelo construido. Finalmente en la Etapa 5 se incluye el estudio y práctica en la herramienta usada.

3.3 Discusión

En la Tabla 1, se puede observar que la etapa más costosa en tiempo es la Etapa 1, Entrevistas.

El principal factor es la documentación de un sistema legado absolutamente desconocido sin documentación previa alguna y con tiempos de dedicación muy acotados de las personas que poseían el conocimiento del mismo.

Por estos motivos dichas entrevistas se debieron hacer lo suficientemente espaciadas y con determinada duración como para no perjudicar el desempeño normal del centro de cómputos del sistema legado que cuenta con un escaso número de funcionarios.

Una alternativa útil y complementaria a las entrevistas con los responsables de los sistemas legados es analizar los datos de los depósitos buscando obtener semántica y cardinalidad de tablas (por ejemplo haciendo uso de un lenguaje de consultas como sql, si el sistema lo soporta).. En nuestro caso particular, esta alternativa pudo utilizarse una vez avanzado el trabajo dado que los datos iniciales se trataban de datos de prueba y las cardinalidades no correspondían a valores razonables. No obstante, una vez disponible los datos reales, si interrogación permitió disminuir las entrevistas.

Tabla 1 – Duración de las Etapas

Etapas			Duración
Etapa 1 Entrevistas	Entrevistas con Nivel Gerencial	20 hs / 2 sems	120 hs / 3 meses y 2 sems
	Entrevistas con Personal S. Legado	100 hs / 12 sems	
Etapa 2 Análisis	Construcción de Diseño Conceptual	20 hs / 3 sems	20 hs / 3 sems
Etapa 3 Diseño Conceptual Multidimensional	Estudio del modelo CMDM	10 hs / 2 días	50 hs / 5 sems y 2 días
	Construcción del modelo y tabla de correspondencias	40 hs / 5 sems	
Etapa 4 Diseño Lógico Multidimensional	Definición de Arquitectura	16 hs / 1 sem	44 hs / 3 sems y 2 días
	Definición del tipo de almacenamiento	10 hs / 1 sem	
	Mapeo del diseño lógico	10 hs / 1 sem	
	Diseño de los procesos de carga y refresque	10 hs / 2 días	
Etapa 5 Implementación	Implementación	200 hs / 6 sems	200 hs / 1 mes y 2 sems

Resulta de gran ayuda en el momento de realizar las entrevistas con el nivel gerencial, explicar de forma somera pero clara, el objetivo de un trabajo de estas características y el tipo de herramienta que se obtendría luego de la implementación. Con ese concepto claro, el usuario de nivel gerencial rápidamente determinó como podrían ser abordadas sus necesidades de información y colaboró de una manera altamente eficaz en la determinación de los requerimientos.

Una sugerencia útil al encarar un trabajo de esta naturaleza es no subestimar el tiempo que puede llevar adquirir el conocimiento necesario de los sistemas legados como para estar en condiciones de reconocer en ellos la información buscada para alimentar el data warehouse.

Conclusiones

Se presenta una metodología para la construcción de una base de datos para ayuda a la toma de decisiones en el área de gestión administrativa del Hospital de Clínicas del Uruguay. Más allá de la metodología en si, este trabajo

permite mostrar las actividades y tiempos que involucra la construcción de un sistema de información con características de apoyo a la toma de decisiones.

El sistema construido se basa en el uso de tecnología conocida en ingles como data warehousing y Olap.

Para el Hospital de Clínicas este trabajo tuvo los siguientes aportes:

- Documentación completa del sistema legado de partida.
- Documentación de diseño del sistema de data warehouse e implementación de un prototipo del cubo de Compras.
- Conocimiento a nivel gerencial de la situación actual de la institución respecto de los sistemas informáticos que necesitan finalización para poder ser integrados a un data warehouse corporativo.
- Conocimientos a nivel gerencial de las posibilidades de apoyo que un data warehouse puede brindar a su gestión.

Si bien se cuenta con mejor tecnología para resolver un sistema de información que ayude a la toma de decisiones, es importante resaltar el tiempo que consume la actividad de preparación para el uso de la misma, en particular el relevamiento de sistemas legados con documentación escasa o inexistente. También se pudo experimentar un alto costo en tiempo debido a la brecha que puede existir entre los requerimientos a nivel gerencial y los sistemas realmente existentes en una institución de gran porte.

Como trabajo futuro mencionamos el relevamiento de otros servicios del Hospital como por ejemplo, CTI, análisis clínicos e ingresos y egresos de pacientes para la inclusión en el data warehouse corporativo de la información y posibilitar resolver requerimientos de toma de decisión en dichos servicios.

La metodología presentada parte de la base que la información de los sistemas legados tiene un nivel de calidad aceptable (por ejemplo, no hay que deducir o identificar conceptos similares a partir de las instancias, todo tiene identificadores claros.) Introducir actividades que resuelvan dicha problemática es además un trabajo futuro.

Finalmente, en este trabajo se consideró la propagación de cambios de datos de los sistemas legados al data warehouse basado específicamente en los sistemas del Hospital de Clínicas. Nos interesa incursionar en el uso de un enfoque como el propuesto por [12] que permita generalizar dicho proceso y a su vez provea facilidades de propagación de cambios a nivel de esquema de los sistema legados hacia el esquema del data warehouse.

Agradecimientos

Agradecemos a todas aquellas personas que colaboraron para que fuese posible la realización de este trabajo, en particular al Dr. Víctor Tonto, A/P Jorge Portugal, y Dr. Daniel Alonso

del Hospital de Clínicas que ofrecieron un tiempo invaluable para responder a nuestras abundantes preguntas y al Msc. Ing. Fernando Carpani del Instituto de Computación por sus discusiones y sugerencias para la aplicación del modelo CMDM.

Referencias

- [1] Chaudhuri S. Dayal U., *An Overview of Data Warehousing and OLAP Technology*. SIGMOD Record, 26(1):65-74, Marzo 1997.
- [2] Revello B., Silva L., *Construcción de un sistema de apoyo a la toma de decisiones para el área gerencial del Hospital de Clínicas*. Informe final proyecto de grado, Facultad de Ingeniería, Universidad de La República, Uruguay, Mayo 2001.
Accesible por [http usando la url http://www.fing.edu.uy/~gutierre/proyectosgrado](http://www.fing.edu.uy/~gutierre/proyectosgrado)
- [3] Kimball R., Reeves L., Ross M., Thornthwaite W.. *The Data Warehouse Lifecycle Toolkit*, John Wiley & Sons, Inc., 1998.
- [4] Thomsen E.. *OLAP solutions. Building Multidimensional Information Systems*. John Wiley & Sons, Inc., 1997.
- [5] Carpani F., Ruggia R.. *An Integrity Constraints Language for a Conceptual Multidimensional Data Model*. XIII International Conference on Software Engineering & Knowledge Engineering. SEKE 2001. Bs. As. Argentina
- [6] Sapia C., Blaschka M., Höfling G., Dinter B.. *Extending the E/R Model for the Multidimensional Paradigm*. Advances in Database Technologies, LNCS Vol. 1552, Springer Verlag, 1999.
- [7] Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P., *Fundamentals of Data Warehouses*. Springer, 2000.
- [8] Inmon W. H.. *Building the Data Warehouse*, John Wiley & Sons, Inc., 1996.
- [9] Inmon W.H.; Welch J.D, Glassey K.L.. *Managing the Data Warehouse*, John Wiley & Sons, Inc., 1996.
- [10] Kimball R., *The Data Warehouse Toolkit*, John Wiley & Sons, Inc., 1996.
- [11] *Microsoft SQL Server 7.0. Resource Guide*, Microsoft Press.
- [12] Marotta A.. *Data Warehouse Design and Maintenance through Schema Transformations*. Master Thesis. October 2000. InCo - Pedeciba, Facultad de Ingeniería, UdelAR, Montevideo, Uruguay. Technical Report INCO 01-01. ISSN: 0797-6410.
Accesible por [http usando la url http://www.fing.edu.uy/~csi/Publicaciones/lista_pub_csi_2000.html](http://www.fing.edu.uy/~csi/Publicaciones/lista_pub_csi_2000.html)

Domicilio para correspondencia

Dirección postal:

Alejandro Gutiérrez

Instituto de Computación, Facultad de Ingeniería, UdelAR

Julio Herrera y Reissig 565 –Piso 5

(11300) Montevideo - URUGUAY

E-mail: gutierre@fing.edu.uy

URL: <http://www.fing.edu.uy/~gutierre>