

Proceso iterativo de diseño del modelo multidimensional de soporte a datos clínicos de pacientes diabéticos*¹

[Iterative design process for multidimensional model to supports clinical data of diabetic patients]

ERNESTO MANGIA², OMAR DENIS³, MA. EMILIA LLORENTE², JAVIER BESSO², ALDO SIGURA², BARTOLOMÉ DROZDOWICZ², ALEJANDRO HADAD²

Recibo: 20.02.2015 – Aprobación: 11.09.2015

Resumen

El presente trabajo propone una estrategia de diseño y desarrollo de un Datawarehouse (DW), para el apoyo a la toma de decisiones de profesionales médicos que atienden pacientes diabéticos. Por pacientes diabéticos, para este desarrollo, se considera a aquellos afectados por diabetes mellitas del tipo 2, por su alta prevalencia y consecuente disponibilidad de datos. Teniendo en cuenta la diversidad de las fuentes de información y la variabilidad de las características y condiciones de los pacientes diabéticos se utilizó un proceso iterativo, basado en los Requerimientos y en las Fuentes de Información existentes. De tal manera se obtuvo una primera versión del Modelo Multidimensional del DW basada en la Historia Clínica para pacientes diabéticos. A partir de dicha versión y para formalizar el proceso de análisis y diseño iterativo

* Modelo para la citación de este artículo:

MANGIA, Ernesto; DENIS, Omar; LLORENTE, Ma. Emilia; BESSO, Javier; SIGURA, Aldo; DROZDOWICZ, Bartolomé & HADAD, Javier (2015). Proceso iterativo de diseño del modelo multidimensional de soporte a datos clínicos de pacientes diabéticos. En: Ventana Informática No. 33 (jul-dic). Manizales (Colombia): Facultad de Ciencias e Ingeniería, Universidad de Manizales. p. 25-38. ISSN: 0123-9678

- 1 Reporte de caso proveniente del Proyecto de Investigación y Desarrollo Plurianual (PIDP) Sistema de Soporte a la Toma de Decisiones basado en datawarehouse para pacientes diabéticos desarrollado en la Facultad de Ciencia y Tecnología de la Universidad Autónoma de Entre Ríos.
- 2 Docentes. Facultad de Ciencia y Tecnología, Universidad Autónoma de Entre Ríos (Oro Verde, Entre Ríos, Argentina). Correos electrónicos: ernesto_mangia@hotmail.com, mellorente@arnet.com.ar, javier@lambdasi.com.ar, aldo.sigura@gmail.com, bdrozdo@santafe-conicet.gov.ar y hadad@santafe-conicet.gov.ar, respectivamente.
- 3 Estudiante avanzado, Licenciatura en Sistemas de Información, Facultad de Ciencia y Tecnología, Universidad Autónoma de Entre Ríos. (Oro Verde, Entre Ríos, Argentina). Correo electrónico: denisomar@hotmail.com.ar

del Modelo Multidimensional, se crean Casos de Uso (CU) para resumir los escenarios operacionales que muestran a los usuarios médicos utilizando el sistema para completar ciertas tareas. Se presentan tres CU los cuales permiten mostrar dos Ciclos de Vida de este proceso iterativo, con la consecuente evolución del Modelo Multidimensional.

Palabras Claves: *Datawarehouse, Modelo Multidimensional, proceso iterativo, Casos de Uso, Fuentes de Información, Diabetes Mellitus tipo 2.*

Abstract

This paper proposes a strategy of design and development of a data warehouse (DW), to support decision making for medical professionals who treat diabetic patients. Given the diversity of information sources, features variability and diabetic patients conditions an iterative process was proposed. This is based on the requirements and existing sources of information used. From first version of the Multidimensional Model (MM) of DW based on the clinical history for diabetic patients was obtained. From this version and to formalize the analysis process and iterative design the MM, Use Cases (UC) was created. This UC summarize the operational scenarios showing medical users using the system to complete certain tasks. Three CU are showing which allow display two Lifecycles of this iterative process, with the consequent evolution of Multidimensional Model.

Keywords: *Datawarehouse, Multidimensional Model, iterative process, Use Case, Information Sources, Diabetes Mellitus type 2.*

Introducción

Los pacientes diabéticos, requieren de un control y seguimiento continuo mediante valoraciones periódicas realizados por profesionales médicos, la toma de datos diarios por el propio paciente en su domicilio y estudios periódicos de laboratorio (estado estable). Además, por ser propensos a complicaciones médicas (estado inestable), como problemas cardiacos, de cicatrización, oftalmológicos, etc., que pueden generar eventos críticos lo que lleva a hospitalización o uso de servicios de tercer nivel de atención como cuidado crítico.

Tales situaciones originan considerable cantidad de datos e información distribuida en variadas escalas de tiempo, generalmente ubicada en diferentes ámbitos y formas, lo que genera dificultades para una

interpretación integral de la evolución de los pacientes, esto indica la posibilidad de integrar y procesar la información relacionada con los distintos eventos, mejorando las decisiones asociadas a diagnósticos y tratamientos.

«En general la necesidad de lograr una información integrada se ha convertido en una prioridad para los niveles de la toma de decisiones de una organización o proceso. La tecnología acompaña estos requerimientos, por un lado a través de las Bases de Datos (BD) las cuales han permitido una administración más segura y ágil de la información. Por otro lado mediante los DW que permiten satisfacer necesidades particulares, facilitando la integración de la información que generalmente procede de diferentes fuentes de datos» (Llorente et al, 2012).

De acuerdo con Elmasri & Navathe (2002), el mejoramiento en las características de herramientas y técnicas analíticas ha conllevado a la posibilidad de almacenar y procesar datos en DW en forma diferente a lo que sucede con las BD. Es de aclarar que la tecnología de DW permite reunir, en un repositorio centralizado, información histórica de un proceso, y que la información de un DW cambia con menor frecuencia y se considera como tiempo no real con actualización periódica.

Coinciden Mazón, Lechtenbörger & Trujillo (2009), Lyman, Scully & Harrison (2008) y Métais et al. (1997), en que las transacciones constituyen la unidad y el agente de cambio de la base de datos de los sistemas transaccionales. No obstante, la información del DW es menos precisa (de grano grueso) y se actualiza de acuerdo a una política elegida, generalmente incremental, realizada por el componente de adquisición del almacén que proporciona todo el procesamiento previo necesario.

Las tareas de modelado para problemática dentro de ámbito médico, en la mayoría de casos, requiere el preprocesamiento de datos para contextualizar el contenido y facilitar la interpretación para la toma de decisiones, que puede involucrar la aplicación de técnicas para abstracción temporal, generación de índices, extracción de características relevantes, identificación de estados, etc., según afirman Hadad et al. (2007), Salvatelli et al. (2007), Durango, Bizai & Drozdowicz (2009) y Evin et al. (2010).

En cuanto a los Niveles de Análisis, se consideró que un DW clínico debe dar soporte al análisis de datos en varios niveles:

- Para este desarrollo el más bajo es el nivel del paciente individual, donde los datos sobre el mismo se pueden visualizar y analizar, por

ejemplo, para encontrar un patrón en el desarrollo de una enfermedad vinculado al mismo.

- El siguiente es el nivel de grupo de pacientes, donde los datos sobre el mismo son analizados, por ejemplo, cuando tengan una enfermedad particular asociada.
- Un nivel más amplio podría ilustrarse a través del caso de una empresa/institución/organización gubernamental de salud, donde clínicos, administradores y especialistas en epidemiología, combinan datos para investigar la calidad, efectividad y eficiencia global de los servicios proporcionados.

Considerando lo anteriormente planteado, el presente trabajo propone una estrategia de diseño y desarrollo de un *Datawarehouse* (DW), para el apoyo a la toma de decisiones de profesionales médicos que atienden pacientes diabéticos.

1. Antecedentes

En la última década se han reportado diferentes enfoques en lo que refiere a metodologías de diseño de *datawarehouse* en el ámbito de la salud. Algunos ejemplos relacionados con estos enfoques son:

- Wisniewski et al. (2003) destacan las dificultades en los pasos para la captura de datos a nivel organizacional en relación al problema del monitoreo de enfermedades infecciosas en un ambiente clínico.
- Zhou et al. (2010) hacen hincapié en el aspecto temporal de largo plazo en relación a los requerimientos y en el rol de los sistemas de información clínicos operacionales con los módulos ETL⁴ en el proceso de diseño.
- Sahama & Croll (2007) remarcen el rol de las etapas de extracción de requerimiento, diseño del modelo de datos, implementación y despliegue, profundizando finalmente en los aspectos de arquitectura, al que le da un rol más relevante para contextos clínicos.
- Ledbetter & Morgan (2001) destacan el rol del proceso de integración de la historia clínica electrónica presente en los sistemas operacionales con un sistema de información clínico basado en *datawarehouse*. En lo que refiere a experiencias más específica relacionada con el diseño de orientado a la gestión de información relacionada con pacientes diabéticos.

4 Extract-Transform-Load (Extraer, Transformar y Cargar)

- Pedersen & Jensen (1998) presentan una caracterización del contexto clínico para diabetes y remarca la dinámica con que se presentan los requerimientos, lo cual tiene un impacto importante en el proceso de diseño.
- Chen et al. (2014) plantean el desarrollo de un sistema basado en *datawarehouse* para diabetes en el cual se enfoca el diseño centrado en los datos del paciente en lugar de abordarlos desde la perspectiva de las enfermedades y tratamientos.

2. Metodología

En el proyecto se consideran las siguientes fuentes de información: Historias Clínicas (Consultorio, Eventos, Medicamentos), Datos de Laboratorio (Tipos de análisis y/o estudios), Datos Recogidos por el paciente en su casa, Base de Datos de Imágenes, Internaciones (Sala Común, Terapia Intensiva (ICU), Ingreso, Eventos, Tratamientos, Medicamentos, Análisis, Egreso), Cirugía, Base de datos de Terapia Intensiva y Base de Datos de Farmacia.

La evolución de las metodologías de diseño de los modelos multidimensionales, planteada por Romero & Abelló (2009), sirve de base para el desarrollo del modelo, identificando los aspectos fundamentales de las metodologías: Requerimientos y Fuentes de información.

A partir de la afirmación de Deo, Deobagkar & Deobagkar (2005), en la cual señalan que si no se tienen perfectamente definidos como se van a analizar los datos, resulta casi imposible comenzar su diseño conociendo todos los requerimientos a priori, considerando la variabilidad de la información y sus fuentes, se llega a definir un proceso de diseño iterativo diferente al proceso incremental, aun cuando este último sea el más convencional.

Crear un modelo lógico de DW incluye estructuras temporales vinculadas a los niveles de las jerarquías dimensionales, que posibiliten el registro de los datos y la recuperación de la información en el tiempo. En el diseño se analiza la correspondencia entre el modelo lógico con las tablas, los atributos de las bases fuentes y los elementos del esquema conceptual, para después analizar la estructura del modelo elegido y extraer la información para plantear diferentes ideas de atributos del MM.

El DW está diseñado asumiendo la conexión con un sistema PACS de almacenamiento de imágenes, donde la comunicación en ambientes de red es la parte medular para el diseño de aplicaciones y se apoya en el protocolo DICOM para la gestión de la imagen diagnóstica.

«– Al final de cada ciclo se entrega una versión completa del software mejorada respecto a la anterior.

– Los ciclos se repiten hasta obtener un producto satisfactorio.

– Los usuarios deben evaluar el producto en cada iteración.

– Se suele aplicar en desarrollos en los que los requisitos no están claros, las primeras versiones pueden ser prototipos, que generalmente son descartados» (Llorente et al, 2012).

Para la implementación del DW se decidió utilizar la plataforma *InfoSphere Platform* (IBM, u.d.), que provee un amplio conjunto de herramientas para brindar soporte a la extracción, limpieza, transformación, carga, almacén y análisis de datos, entre otros.

El proceso realizado para la implementación de *Dim_Paciente* y *Patient*, se crea un proyecto de tipo *Analyzerproject* en *DataStage Administrator*, que contiene las definiciones de las fuentes de información, trabajos (procesos ETL) y metadatos, así como un Nombre de Origen de Datos (DSN) para la base de datos de Historia Clínica y otro para el DW (con DB2 como motor de base de datos). Luego, se crean, en *DataStage Designer*, los *Data Connection Object* (DCO) utilizados para vincular los DNS definidos, se importa la definición de la FI (metadatos) y se importan los metadatos correspondientes a las tablas del DW.

Se diseñaron las ETL utilizando el módulo *DataStage Designer*, así como dos objetos *ODBC Connector* para la BD de HC y la DW, los cuales se unieron mediante un proceso *Transformer* para transformar los datos de origen y almacenarlos en el destino, donde se utiliza la función *Split* para dividir el campo *P_NAME* del origen en *NOMBRE* y *APELLIDO* en el destino. Por último, se ejecuta la ETL, con un total de 100 registros de la tabla *PATIENT*, obteniendo como resultado 100 registros en el destino transformados (Figura 1).

3. Resultados

Para generar este proceso iterativo basado en Casos de Uso se propone una primera versión del Modelo Multidimensional basado en una parte importante de las fuentes información, como es la Historia Clínica para pacientes diabéticos.

La primera definición de las posibles entidades relevantes, con sus atributos y su clave primaria, se resume así: Diabetes: (*id_Diabetes*, *Tipo*; *Descripción*), *Diabet*; Des (*Id_medico*, *Nombre_Apellido*, *Telefono*, *Direccion*), *Tratamiento*: (*Id_Tratamiento*, *Descripcinto*), *Medicacion*

(*Id_Medicacion, Nombre, Dosis*), *Estilo de Vida* (*Id_EV, Dieta, Actividad_fisica, Sobrepeso, Adicciones, Situacipeso*), *Diagnostico 2rio* (*Id_Ds, Tipo, Descripcion*), *Resultado de Laboratorio 2rio* (*Id_RI2, Fecha, Hemograma, Hepatograma, Orina, Otros_Analisis*), *Profesional* (*Id_Medico, Especialidad, Nombre_Apellido, Telefono, Direccion*), *Tratamiento 2rio* (*Id_T2, Descripci 2r*), *Medicacion 2ria* (*Id_Medicacion, Nombre, Dosis*) y *Diagnostico por imticoosis* (*Id_DI, Fecha, Descripciripcino, Direc*).

The screenshot shows the 'Control Center - DB2COPY' application. The top right pane displays a data flow diagram with components: 'Clinical' (100 rows, 50 rows/sec), 'Split', 'pacientes' (100 rows, 50 rows/sec), and 'DW'. The bottom pane shows a table with the following data:

NOMBRE	APELLIDO	SEXO	TELEFONO	ID	DIRECCION	FECHA_NAC
Raquel	Carbajo Escobano	F		02532ADC-20...		29-ene-1968
Pilar	Madrid Vergara	F		105156FE-1A0...		20-abr-1978
Manuel	García Garrido	M		118C44F6-C1E...		25-abr-1965
José Manuel	Saura Regueira	M		1497B069-484...		25-jun-1951
Antonio	Benavides Ares	M		157F512A-F32...		25-mar-1940
Pablo	García Rivero	M		188B591A-6E...		10-feb-1967
Antonio	González Mendizábal	M		1A0B8B72-56...		23-jul-1972
Carmen	Campos Esteve	F		22A24470-55F...		03-jun-1969
Natalia	Martínez Fernández	F		22B289F1-C1F...		25-nov-1965
Rafael	Muñoz Granados	M		4E0C26DE-FFE...		04-sep-1974
María Isabel	Egea Sánchez	F		4F584A79-A6...		08-jul-1993

Figura 1. Tabla Dim_Paciente como consecuencia de la ETL Paciente ejecutada exitosamente (tiene una ventana superpuesta en el extremo superior derecho)

Al analizar las relaciones entre las entidades definidas y la relevancia de ellas, y estudiar el aporte de cada una a la resolución de los requerimientos definidos, con el objetivo de proponer un Modelo Multidimensional, de acuerdo con Llorente et al. (2013) «es posible definir que las mismas están basadas en dos aspectos fundamentales: *Requerimientos y Fuentes de Información existentes*», se definen las dimensiones del modelo (Figura 2).

La definición de las tablas de hechos se consideraron el diagnóstico base (orientado a registrar la evolución de la enfermedad) y el diagnóstico secundario (registro de eventos clínicos relacionados o no con la enfermedad base), para luego relacionar las tablas, definiendo las claves primarias y foráneas y obtener el primer Modelo Multidimensional propuesto (Figura 3). Para formalizar el proceso de análisis y diseño

iterativo del Modelo Multidimensional, se crean *Casos de Uso (CU)* para resumir los escenarios operacionales, que muestran a los usuarios médicos utilizando el sistema para completar ciertas tareas, y desarrollar servicios computacionales que soporten los mismos.

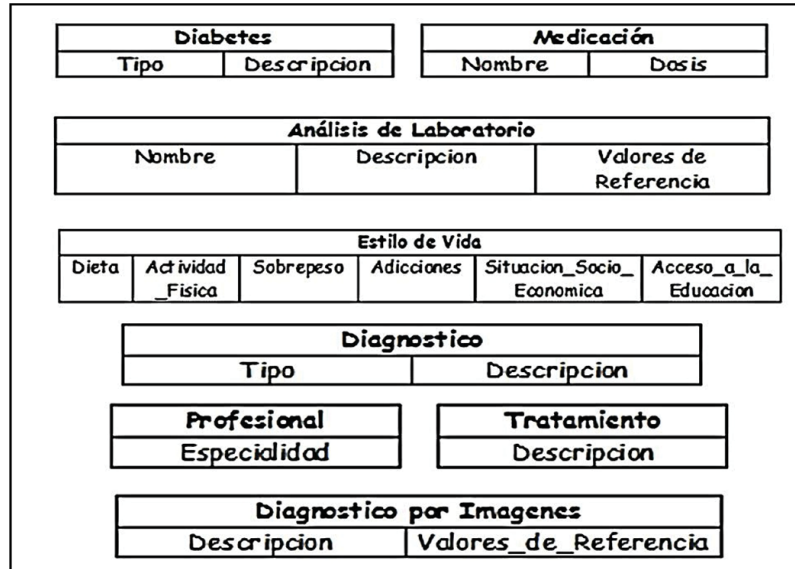


Figura 2. Dimensiones del Modelo (Llorente et al., 2014, 3)

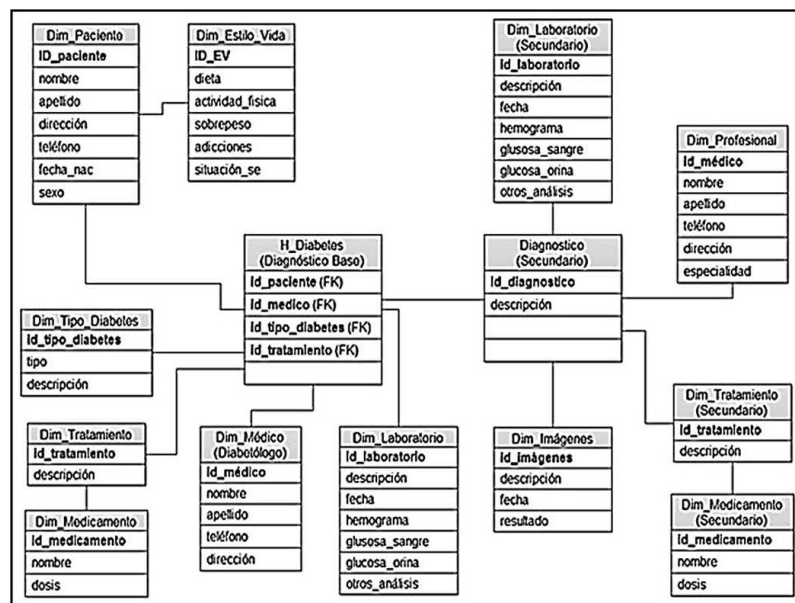


Figura 3. Primer Modelo Multidimensional propuesto (Llorente et al., 2014, 3)

Debido a que solo algunos requerimientos de proceso de la información pueden conocerse antes de comenzar, debe utilizarse una forma de desarrollo del DW que difiera del ciclo de vida tradicional, por lo que se propone un Ciclo de Vida Iterativo, con repetición de varios ciclos de vida en cascada. En él se consideran dos criterios: - obtener un prototipo rápidamente para ir validando las diferentes partes generadas para dar soporte a los casos de uso, y - durante la incorporación de nuevos CU no generar nuevas funciones o que los pocos que se generen no modifiquen significativamente el modelo lógico.

En la Tabla 1, se describe un Ciclo del proceso iterativo propuesto, con tres CU, que corresponden a tareas de diferente complejidad realizadas por el actor considerado en primera instancia (el médico), desde el rol habitual de atención al paciente, hasta un análisis que involucre grupos de pacientes centrándose en el análisis de, por ejemplo, una patología de interés.

Continuando con el diseño iterativo del Modelo Multidimensional, en este caso basado en Casos de Uso, se realiza un análisis de la estructura de datos seleccionada como fuente de información y la estructura del modelo propuesto, para obtener un nuevo modelo modificado (Figura 4).

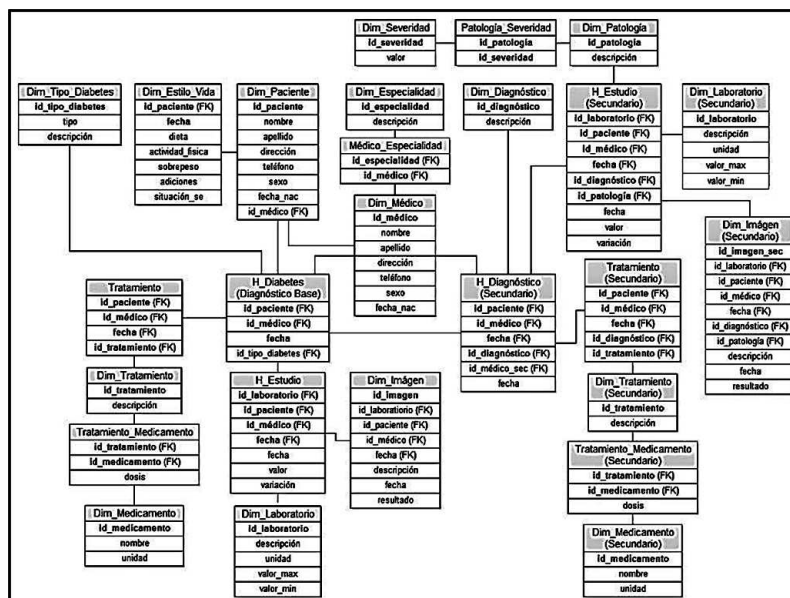


Figura 4. Modelo Multidimensional modificado

Tabla 1. Casos de uso

Caso de uso	Recuperación de Imágenes	Comparación de Imágenes	Análisis de un Grupo de Pacientes
Objetivo	Recuperar dos imágenes realizadas al mismo paciente en tiempos diferentes	Comparar imágenes para detectar diferencias en alguna evolución de la patología y mostrar al usuario las imágenes crudas y los resultados de la comparación	Estadística para un grupo de pacientes que comparten condiciones patológicas equivalentes
Contexto	La identidad del médico y sus privilegios de acceso han sido validados	La identidad del médico y sus privilegios de acceso han sido validados	La identidad del médico y sus privilegios de acceso han sido validados
Actores	Médico	Médico	Médico
Recursos	El PACS y el DataWarehouse están accesibles	El PACS y el DataWarehouse están accesibles	El PACS y el DataWarehouse están accesibles
Episodios	Acción del Actor: Responsabilidades del sistema a. El médico ingresa el nombre del paciente (o No. de registro): Listar los estudios según el paciente b. El médico elige el tipo de estudio: Listar los estudios según el paciente y el tipo de estudio, indicando el momento de realización de cada uno c. El médico selecciona una o más imágenes: Transferir las imágenes a la computadora cliente	Acción del Actor: Responsabilidades del sistema a. El médico ingresa el nombre del paciente (o No. de registro): Listar los estudios según el paciente. b. El médico elige el tipo de estudio Listar los estudios según el paciente y el tipo de estudio, indicando el momento de realización de cada uno c. El médico selecciona dos o más imágenes: Transferir las imágenes a la computadora cliente d El médico elige dos imágenes para comparar: Ejecutar el método de comparación, Generar Imagen resultado de la comparación y Transferir imagen resultado a la computadora cliente	Acción del Actor: Responsabilidades del sistema a. El médico ingresa el nombre o selecciona una patología nombrada según una clasificación estandarizada (CIE9, CIE10, SNOMED, etc.): Listar los estudios asociados a ella indicando la severidad de la enfermedad diabética al momento del último estudio b. El médico elige el grupo de pacientes diabéticos con mayor nivel de severidad: Listar los estudios según los pacientes pertenecientes a este grupo c. El médico selecciona, dentro del último grupo, los estudios realizados cuando el nivel de severidad diferente: Listar los estudios según el nivel de severidad d. El médico busca los estudios más frecuentes según el nivel de severidad elegido: Ejecutar el método estadístico estándar, y Transferir resultado a la computadora cliente

Las modificaciones básicas realizadas son: - la granularidad del *H_Diagnóstico* (Base), - unificación de las dimensiones Profesional y Médico, - adición de la dimensión Especialidad, - adición de *H_Estudio*, - se agregan valores de referencia en la dimensión Laboratorio, - soporte (un tratamiento puede requerir muchos medicamentos y un medicamento pueden ser utilizado en varios tratamientos), - se separan los estudios y laboratorios secundarios de los asociados al diagnóstico base, - *H_Diabetes* almacena información de los eventos o consultas en una determinada fecha y - la consulta puede tener asociados muchos tratamientos.

Luego se define un plan de tareas para aplicar, que comprende - Seleccionar al menos dos nuevos casos de uso, - Analizar la satisfacción de los requerimientos de los CU seleccionados, - Analizar el impacto en el modelo definido, - Proponer un nuevo modelo de DW que satisfaga los nuevos casos de uso, - Implementar una nueva fuente, si es necesaria - Implementar las modificaciones sobre el modelo de DW, y - Definir las ETL de acuerdo con las modificaciones realizadas. Además, se analiza la reconversión de la información ya transformada y cargada en la primera iteración, para proponer el nuevo modelo mostrado en la Figura 5.

A partir del primer modelo propuesto como solución, se determina la relación entre la fuente de información seleccionada y dicho modelo, que implica definir ETL de fuentes de información de distinto tipo (de datos y de imágenes y señales). La FI propuesta se encuentra soportada en el motor de base de datos DB2, también proporcionado por IBM.

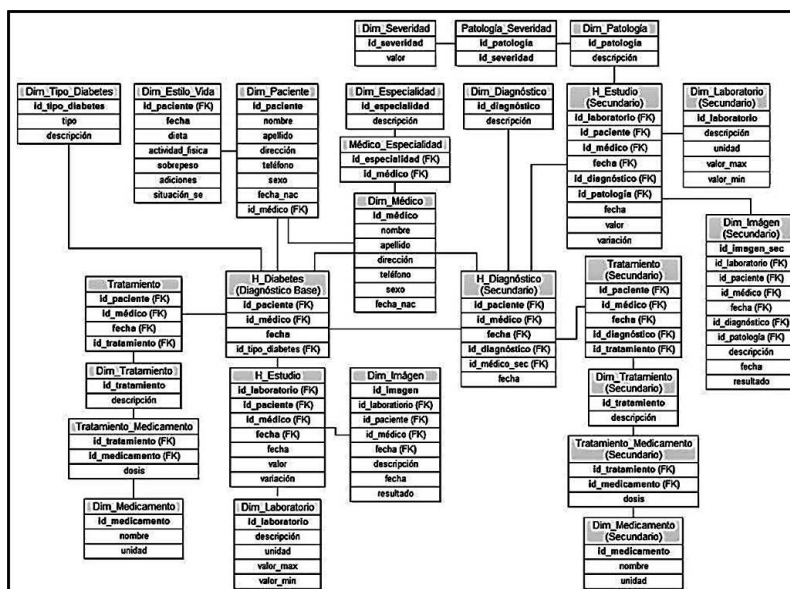


Figura 5. Modelo Multidimensional segundo Ciclo de Vida

4. Conclusiones

Se propone un proceso iterativo para el diseño del Modelo Multidimensional del DW, teniendo en cuenta la diversidad de las fuentes de información y la variabilidad de las características y condiciones de los pacientes diabéticos. Esto genera que generalmente los usuarios médicos de un DW clínico, como el propuesto en este trabajo, no tienen perfectamente definido la manera de analizar los datos y por lo tanto resulta casi imposible comenzar su diseño conociendo todos los requerimientos a priori.

Una primera versión del Modelo Multidimensional se basa en una parte central de las fuentes información, la Historia Clínica para pacientes diabéticos. En ella se analiza la estructura del modelo elegido y extrae la información para plantear las posibles entidades relevantes, con sus atributos y su clave primaria del Modelo Multidimensional. Se analizan, además, las relaciones existentes entre las entidades definidas y la relevancia de cada una de ellas, estudiándose su aporte en la resolución de los requerimientos.

Del análisis realizado, se definen las dimensiones y las tablas de hechos. Para formalizar el proceso de análisis y diseño iterativo del Modelo Multidimensional, se crean Casos de Uso (CU) para resumir los escenarios operacionales que muestran a los usuarios médicos utilizando el sistema y se proponen las características de los Ciclos de Vida que conforman el proceso iterativo de diseño.

La aplicación del primer CU completa un Ciclo de Vida del proceso iterativo y genera una nueva versión del Modelo Multidimensional. Por su parte el segundo Ciclo de Vida del proceso con la consecuente versión del modelo Multidimensional se logra con la aplicación de dos CU restantes. En todos los casos, el Modelo Multidimensional resulta ser una evolución del generado en el Ciclo de Vida anterior, lo cual demuestra la factibilidad del proceso iterativo de diseño a través de la aplicación de nuevos CU.

5. Referencias bibliográficas

- CHEN, L.; CHIOU-SHIANG, W.; I-CHING, W.; HUI-CHU, Y.; HUI-YU, P.; HUI-CHUEN, C.; CHIA-HSIUN, C.; TIEN-JYUN, C.; YI-DER, J.; LEE-MING, C. & FEIPEI, L. (2014). Patient-centric Data Warehouse Design: An Empirical Study Applied in Diabetes care [online]. In: The Sixth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies, Biotechno 2014 (20-24/04/2014). Chamonix (France): International Academy, Research and Industry Association, IARIA. p. 43-47. ISBN: 978-1-61208-335-3.
- DEO, S.S.; DEOBAGKAR, D.N. & DEOBAGKAR, D.D. (2005). Design and development of a web-based application for diabetes patient data management. In: Informatics in Primary Care, Vol. 13, No. 1 (feb). London (UK): Radcliffe Publishing Ltd. p. 35-41. ISSN: 1476-0320.

- DURANGO LONDOÑO, N.; BIZAI, G. & DROZDOWICZ, B. (2009). Implementación y aplicación de algoritmos Retinex al preprocesamiento de imágenes de retinografía color [en línea]. En: Revista Ingeniería Biomédica, Vol. 3, No. 6 (jul-dic). Medellín (Colombia): Universidad CES, Escuela de Ingeniería de Antioquia. p. 36-43. ISSN 1909-9762 <<http://repository.eia.edu.co/revistas/index.php/BME/article/view/75/75>> [consulta: 13/02/2015]
- ELMASRI, R. & NAVATHE, S.B. (2002). Fundamentos de Sistemas de Bases de Datos. 3 ed. , Madrid (España): Pearson Educación. 962 p. ISBN: 9788478290512
- EVIN, D.; HADAD, A.; MARTINA, M. & DROZDOWICZ, B. (2011). Predicción de estados de hipotensión empleando Modelos Ocultos de Markov [en línea]. En: Revista Facultad de Ingeniería, Vol. 20, No. 30 (ene-jun). Tunja (Colombia): Universidad Pedagógica y Tecnológica de Colombia. p. 55-63. ISSN: 0121-112 <<http://revistas.uptc.edu.co/revistas/index.php/ingenieria/article/view/1418/1413>> [consulta: 23/05/2015]
- HADAD, A.J.; EVIN, D.A.; DROZDOWICZ, B. & CHIOTTI, O. (2007). Temporal Abstraction for the Analysis of Intensive Care Information [online]. In: Journal of Physics: Conference Series, Vol. 90 [16th Argentine Bioengineering Congress (SABI 2007) and the 5th Conference of Clinical Engineering (26-28/09/2007), San Juan (Argentina): IOP Science]. Bristol (UK): IOP Publishing. p. 1-7. ISSN: 1742-6596 <doi:10.1088/1742-6596/90/1/012074> <http://iopscience.iop.org/1742-6596/90/1/012074/pdf/1742-6596_90_1_012074.pdf> [consult: 25/05/2015]
- IBM (u.d.). Infosphere DataStage [online]. New Orchard Road Armonk (NY, USA): IBM Corporation. <<http://www-03.ibm.com/software/products/en/ibminfodata>> <http://www-947.ibm.com/support/entry/portal/Overview/Software/Information_Management/InfoSphere_DataStage> [consult: 25/05/2015]
- LEDBETTER, C.S. & MORGAN, M.W. (2001). Toward best practice: Leveraging the electronic patient record as a clinical data warehouse [online]. In: Journal of Healthcare Information Management, Vol. 15, No. 2 (Summer). Chicago (Il, USA): of Healthcare Information Management Systems Society, HIMSS. p. 119-131. ISSN: 1099-811X
- LLORENTE, M.E.; SIGURA, A.; BESSO, J.; HADAD, A. & DROZDOWICZ, B. (2012). Proceso de Diseño basado en Casos de Uso para un Datawarehouse Clínico [en línea]. En: XVIII Congreso Argentino de Ciencias de la Computación, CACIC 2012 (08-12/10/2012). Bahía Blanca (Buenos Aires, Argentina): Red de Universidades Nacionales con carreras en Informática, RedUNCI. Anales del XVIII Congreso Argentino de Ciencias de la Computación – IX Workshop de Bases de Datos y Minería de Datos, WBDDM. p. 1029-1038 ISBN: 978-987-1648-34-4. <http://sedici.unlp.edu.ar/bitstream/handle/10915/23745/Documento_completo.pdf?sequence=1> [consulta: 12/02/2015]
- LLORENTE, M.E.; SIGURA, A.; BESSO, J.; MANGIA, E.; HADAD, A.J. & DROZDOWICZ, B. (2013). Análisis de fuentes de información para el proceso de diseño de un datawarehouse sobre pacientes diabéticos [en línea]. En: XV Workshop de Investigadores en Ciencias de la Computación, WICC 2013 (18-19/04/2013). Paraná (Entre Ríos, Argentina): Red de Universidades Nacionales con carreras en Informática, RedUNCI. CALUVA, Claudio; ARANGUREN, Silvia Mónica & MUZACHIODI, Rodolfo (eds.) (2013). WICC 2013. Entre Ríos (Argentina): UADER. p. 102-106. ISBN: 9789872817961. <http://sedici.unlp.edu.ar/bitstream/handle/10915/27106/Documento_completo.pdf?sequence=1> <<http://sedici.unlp.edu.ar/handle/10915/27073>> [consulta: 12/02/2015]
- LLORENTE, M.E.; SIGURA, A.; BESSO, J.; MANGIA, E.; HADAD, A.J.; MANCINI, M.R.; QUIJADA, N. & DROZDOWICZ, B. (2014). Propuesta de un Modelo Multidimensional para un Datawarehouse sobre pacientes diabéticos [en línea]. En: XVI Workshop de Investigadores en Ciencias de la Computación, WICC 2014 (07-08/05/2014). Ushuaia (Tierra del Fuego, Argentina): Red de Universidades Nacionales con carreras en Informática, RedUNCI. p. 173-177. ISBN: 978-950-34-1084-4. <http://sedici.unlp.edu.ar/bitstream/handle/10915/41988/Documento_completo.pdf?sequence=1> <<http://wicc2010.info.unlp.edu.ar/files/WICC2014-articulos-publicados.pdf>> [consulta: 13/02/2015]
- LYMAN, J.A.; SCULLY, K. & HARRISON Jr., J.H. (2008). The Development of Health Care Data Warehouses to Support Data Mining [online]. In: Clinics in Laboratory Medicine, Vol. 28, No. 1 (mar). Amsterdam (NX): Elsevier B.V. p. 55-71. ISSN: 0272-2712 <doi:10.1016/j.cll.2007.10.003> <[http://www.labmed.theclinics.com/article/S0272-2712\(07\)00112-6/fulltext](http://www.labmed.theclinics.com/article/S0272-2712(07)00112-6/fulltext)> [consult: 12/02/2015]

- MAZÓN, J.N.; LECHTENBÖRGER, J. & TRUJILLO, J. (2009). A survey on summarizability issues in multidimensional modeling [online]. In: *Data & Knowledge Engineering*, Vol. 68, No. 12 (dec). Cambridge (MA, USA): Elsevier B.V. p. 1452-1469. ISSN: 0169-023X. <doi:10.1016/j.datak.2009.07.010> <http://www.sciencedirect.com/science/article/pii/S0169023X09001104> [consult: 13/02/2015]
- MÉTAIS, E.; KEDAD, Z.; COMYN-WATTIAU, I. & BOUZEGHOUB, M. (1997). Using linguistic knowledge in view integration: Toward a third generation of tools. *Data & Knowledge Engineering*, Vol. 23, No. 1 (jun). Cambridge (MA, USA): Elsevier B.V. p. 59-78. ISSN: 0169-023X <doi:10.1016/S0169-023X(97)00004-9> <http://www.sciencedirect.com/science/article/pii/S0169023X97000049> [consult: 12/02/2015]
- PEDERSEN, T.B. & JENSEN, C.S. (1998). Research Issues in Clinical Data Warehousing [online]. In: 10th International Conference on Scientific and Statistical Database Management, of SSDBM'98 (01-03/07/1998). Capri (Italy): IEEE. Proceedings of the SSDBM'98. p. 43-52. ISBN: 0-8186-8575-1. <doi:10.1109/SSDM.1998.688110> <http://people.cs.aau.dk/~csj/Papers/Files/1998_pedersenISS.pdf> [consult: 12/02/2015]
- ROMERO, O. & ABELLÓ, A. (2009). A survey of Multidimensional Modelling Methodologies [online]. In: *International Journal of Data Warehousing and Mining*, Vol. 5, No. 2 (apr-jun). Hershey (PA, USA): IGI Publishing p. 1-23. e-ISSN: 1548-3932, ISSN: 1548-3924 <http://www.essi.upc.edu/~aabello/publications/09.IJDWM.pdf> <doi:10.4018/jdwm.2009040101> [consult: 15/02/2015]
- SAHAMA, T.R. & CROLL, P.R. (2007) A Data Warehouse Architecture for Clinical Data Warehousing [online]. In: 5th Australasian symposium on ACSW frontiers, ACSW '07 (30/01-02/02/2007). Ballarat (Australia): Australian Computer Society Inc. BRANKOVIC, Ljiljana; CODDINGTON, Paul; RODDICK, John F.; STEKETEE, Chris; WARREN, James R. & WENDELBOURN, Andrew (eds.) (2007). Proceedings of the fifth Australasian symposium on ACSW Frontiers op on Health Knowledge Management and Discovery (HKMD 2007) CRPIT, Vol. 68, Darlinghurst (Australia): Australian Computer Society, Inc. p. 227-232. ISBN: 1-920-68285-X. <http://dl.acm.org/citation.cfm?id=1274531.1274560> [consult: 15/02/2015]
- SALVATELLI, A.; BIZAI, G.; BARBOSA, G.; DROZDOWICZ, B. & DELRIEUX, C. (2007). A comparative analysis of preprocessing techniques in colour retinal images [online]. In: *Journal of Physics: Conference Series*, Vol. 90 [16th Argentine Bioengineering Congress (SABI 2007) and the 5th Conference of Clinical Engineering (26-28/09/2007), San Juan (Argentina): IOP Science]. Bristol (UK): IOP Publishing. p. 1-7. ISSN: 1742-6596 <doi:10.1088/1742-6596/90/1/012069> <http://iopscience.iop.org/1742-6596/90/1/012069/pdf/1742-6596_90_1_012069.pdf> [consult: 25/05/2015]
- WISNIEWSKI, M.F.; KIESZKOWSKI, P.; ZAGORSKI, B.M.; TRICK, W.E.; SOMMERS, M. & WEINSTEIN, R.A. (2003). Development of a Clinical Data Warehouse for Hospital Infection Control [online]. In: *Journal of the American Medical Informatics Association, JAMIA*, Vol. 10, No. 5 (sep-oct). Bethesda (MD, USA): American Medical Informatics Association, AMIA. p. 454-462. e-ISSN: 1527-974X. <doi:10.1197/jamia.M1299> [consult: 13/02/2015]
- ZHOU, X.; CHEN, S.; LIU, B.; ZHANG, R.; WANG, Y.; LI, P.; GUO, Y.; ZHANG, H.; GAO, Z. & YAN, X. (2010). Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support [online]. In: *Artificial Intelligence in Medicine*, Vol. 48, No. 2-3 (feb-mar). Philadelphia (PA, USA): Elsevier Inc. p. 139-152. ISSN: 0933-3657 <http://dx.doi.org/10.1016/j.artmed.2009.07.012> [consult: 13/02/2015]