

Evolución del Modelo Multidimensional en un DataWarehouse para pacientes diabéticos

E. Mangia¹, D. Omar¹, J.P. Madonni¹, M. E. Llorente¹, J. Besso¹, A. Sigura^{1,2}, A. J. Hadad^{1,2},
B. Drozdowicz^{1,2}

¹Facultad Ciencia y Tecnología, Universidad Autónoma de Entre Ríos

²Facultad Ingeniería, Universidad Nacional de Entre Ríos

Ruta 11, Oro Verde, Entre Ríos, Argentina

mellorente@arnet.com.ar, bdrozdo@santafe-conicet.gov.ar

Resumen

En el presente trabajo se propone una nueva aplicación de la estrategia de diseño iterativo para el Modelo Multidimensional de un DataWarehouse propuesto en etapas anteriores y que respondía al primer caso de uso definido. Este caso de uso estaba orientado a responder consultas sobre la condición de la enfermedad de un paciente. La propuesta actual es mostrar la evolución del modelo aplicando los requerimientos definidos en el segundo caso de uso, orientado a evaluar la enfermedad en grupo de pacientes. Los Casos de Uso (CU) que fueron definidos inicialmente en el trabajo de investigación, representan los diferentes requerimientos planteados por usuarios. Esta forma iterativa de diseño permite la incorporación de nuevos requerimientos desde los distintos ámbitos de interés en el tema.

Palabras clave: Datawarehouse, Pacientes Diabéticos, Casos de Uso, Estructuras de Datos, Modelo Multidimensional, ETL.

Contexto

El presente trabajo se inserta en un Proyecto de Investigación Plurianual (PIDP) denominado “*Sistema de Soporte a la Toma de Decisiones basado en datawarehouse para pacientes diabéticos*”. Dicho proyecto es desarrollado en

la Facultad de Ciencia y Tecnología de la Universidad Autónoma de Entre Ríos (FCYT - UADER).

Introducción

De acuerdo a la propuesta de una Metodología Iterativa de diseño, en cada ciclo, el objetivo es poder resolver requerimientos propuestos [1]. Esto permite ampliar el modelo y permitir dar respuesta a nuevos requerimientos. Fundamentalmente cuando en el problema inicial planteado, no se conocen la totalidad de los requerimientos de los distintos usuarios [2]. El diseño obtenido anteriormente aplicando el concepto del *Modelo Multidimensional* [3], estaba basado en los dos Casos de Uso orientados al nivel de paciente [2]. El objetivo de este trabajo es presentar el análisis necesario para lograr la evolución del mencionado diseño frente a los requerimientos de un tercer *Caso de Uso* [2]. Los requerimientos definidos en los *Casos de Uso*, están divididos en niveles [2]. La evolución se realiza utilizando el Caso de Uso orientado al segundo nivel que corresponde al *Grupo de Pacientes*.

Líneas de investigación y desarrollo

De las líneas de investigación descriptas para

este proyecto en el trabajo presentado en el WICC 2012 [4], en esta propuesta se analizaron las siguientes líneas:

1. Estructuras de datos representativas del dominio de análisis.
2. Métodos ETL para fuentes de información de referencia.
5. Granularidad y dimensiones del modelo.
6. Metadatos y reglas de negocios.

Resultados y Objetivos

El Modelo Multidimensional del DW para enfermos diabéticos que es utilizado para la implementación del prototipo es el representado en la Fig. 2. Este modelo resulta de los requerimientos propuestos en el primer Caso de Uso (Fig. 1)

Recuperación de Imágenes
Objetivo: Recuperar dos imágenes realizadas al mismo paciente en tiempos diferentes.
Contexto: La identidad del médico y sus privilegios de acceso han sido validados
Actores: Médico
Recursos: El PACS y el DataWarehouse están accesibles
Episodios (*):
a. El médico ingresa el nombre del paciente (o nro 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

Figura 1. Casos de Uso orientados al primer nivel

Sobre este primer diseño también se realizó un trabajo de análisis y definición de las ETL's que permiten el correspondiente procesamiento de información [5].

A continuación, se analizó el diseño de la Fig.2 de acuerdo a los requerimientos propuestos en el Caso de Uso orientado al segundo nivel de usuarios (Fig.3).

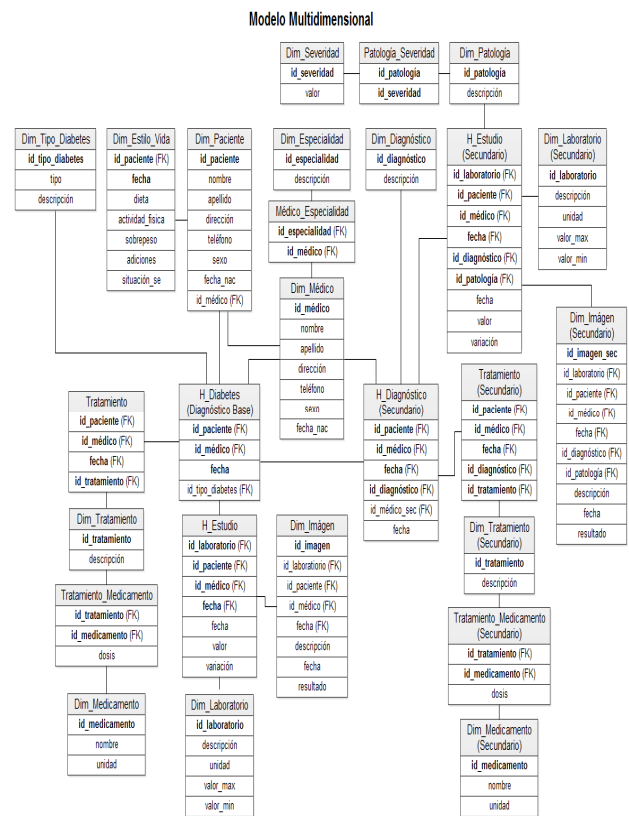


Figura 2. Modelo Multidimensional propuesto.

Analizando los requisitos detallados en ambos Casos de Uso, se observa necesaria la incorporación de información que permita comparar resultados de un análisis con los valores de referencia. Se decide incluir en la fuente de información aquella que permita realizar esta comparación, la cual permita comparar el resultado de estudios de pacientes y establecer el desvío del mismo. De esta forma sería factible relacionar el resultado de un estudio con una patología determinada.

Este requerimiento responde al Caso de Uso correspondiente a la Fig.3 (1er columna).

Teniendo en cuenta a las nuevas necesidades, se se propone el nuevo modelo multidimensional mostrado en la Fig. 4, que es una evolución respecto del anterior.

Comparación de Imágenes	Análisis de un Grupo de Pacientes
Objetivo: Comparar imágenes para detectar diferencias en alguna evolución de la patología y mostrar al usuario las imágenes crudas y los resultados de la comparación	Objetivo: Estadística para un grupo de pacientes que compartan condiciones patológicas equivalentes.
Contexto: La identidad del médico y sus privilegios de acceso han sido validados	Contexto: La identidad del médico y sus privilegios de acceso han sido validados
Actores: Médico	Actores: Médico
Recursos: El PACS y el DataWarehouse están accesibles	Recursos: El PACS y el DataWarehouse están accesibles
Episodios (*): a. El médico ingresa el nombre del paciente (o nro 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	Episodios (*): a. El médico ingresa / selecciona una patología nombrada. → Listar los estudios asociados a dicha patología indicando la severidad diabética b. El médico elige el nivel de severidad que desea observar → Listar todos los estudios según los pacientes pertenecientes a este nivel de severidad. c. El médico selecciona las variables a analizar (edad, peso, sexo, etc) en los estudios → Ejecutar el método estadístico estándar y graficar histograma para representar media, mediana, moda, entre otras. d. El médico selecciona el subgrupo según el nivel de severidad → Transferir Resultado a la computadora cliente.

Figura 3. Caso de Uso orientados al segundo nivel

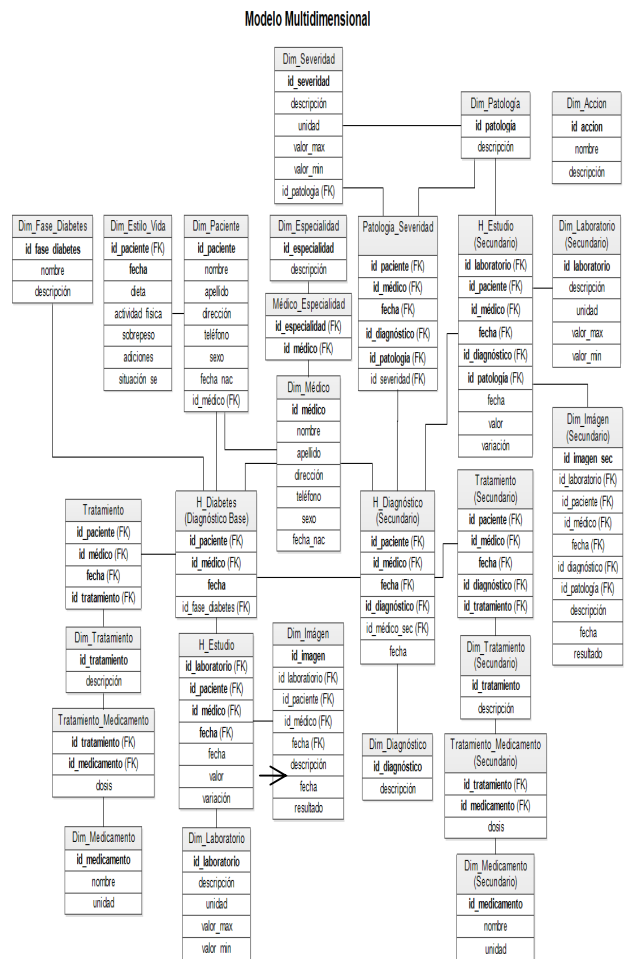


Fig. 4 Modelo Multidimensional para el segundo Ciclo de Vida

Se aplicaron las siguientes modificaciones:

- Se agregó la dimensión Dim_Patología asociada a la H_Estudio
- Esta nueva dimensión se relaciona con la severidad a través de Patología_Severidad que permite definir la dimensión Dim_Severidad
- También se agregó la dimensión Dim_Imagen relacionada con la tabla de hechos H_Estudio (Secundario) entendiendo que también estudios secundarios pueden incluir imágenes como resultados.

Respecto del Caso de Uso representada en la 2da.columna de la Fig.3, las modificaciones

son:

- Se agregaron los campos unidad, valor_max, valor_min a la dimensión Dim_Severidad con el objetivo de almacenar los valores de referencia de la misma.
- Se modificó la cardinalidad “muchos a muchos” entre Dim_Severidad y Dim_Patologia por la cardinalidad “1 a muchos” (una patología puede tener n severidades y una severidad pertenece solo a una patología). Esto es debido a que diferentes severidades pueden tener la misma descripción (por ejemplo: Normal) pero pueden representar diferentes valores de referencia dependiendo de la patología.
- Se modificó Dim_Tipo_Diabetes por Dim_Fase_Diabetes, para determinar el estado de evolución de la diabetes.

De esta manera, se ha obtenido un modelo que responde a diferentes requerimientos y fundamentalmente aplicando el proceso iterativo propuesto inicialmente [6].

Sobre este diseño, no solamente se analizó la factibilidad de dar respuesta a los requerimientos propuestos en los Casos de Uso, sino que además se modificaron las fuentes de información tomadas como origen. Este análisis es fundamental para poder redefinir los procesos de ETL's necesarios para la carga de información.

Se propone como camino a seguir, la aplicación de una nueva iteración aplicando requerimientos orientados a un tercer nivel de usuarios (instituciones y organizaciones médicas) [2].

Esta necesidad, impone la definición de un nuevo Caso de Uso que respete las necesidades de este nuevo nivel de usuarios. A tal efecto, se propone el Caso de Uso de la Fig.5, el cual tiene por objetivo conocer si la cantidad de médicos que actualmente trabaja en la institución es suficiente para atender todas las

consultas de los pacientes.

Análisis de los pacientes atendidos por una institución de salud	
Objetivo	Determinar si la cantidad de médicos de la institución es suficiente para atender todas las consultas de pacientes
Contexto	La identidad del usuario y sus privilegios de acceso han sido validados
Actores	Dirección médica y autoridades institucionales
Recurso	El DW está accesible
Episodios	<i>Acción del actor ⇒ Responsabilidades del sistema El usuario selecciona las variables que pueden combinarse con la presente consulta (especialidad, patología, edad, sexo, peso, estatura, etc) ⇒ El sistema ejecuta la consulta junto con las variables seleccionadas y genera un reporte para el usuario.</i>

Figura 5. Caso de Uso orientados al tercer nivel

Para poder dar respuesta a este nuevo CU, es necesario disponer de información relacionada a los turnos de todos los profesionales médicos, como así también del tiempo de cada consulta de paciente. Con dicha información es posible calcular el promedio de pacientes diarios que atiende cada médico y de esta forma poder determinar si la cantidad de médicos es suficiente para cubrir todas las consultas de los paciente.

El modelo multidimensional propuesto en la Fig. 4 no almacena la información requerida para dar soporte CU anteriormente definido.

Es trabajo futuro realizará una nueva evolución para aplicar las modificaciones necesarias al modelo multidimensional y de esta forma dar respuesta al CU referido al tercer nivel de usuarios.

Formación de Recursos Humanos

El equipo de trabajo está conformado por especialistas del área de sistemas de información, inteligencia artificial y bioingeniería. Integrantes del equipo tienen formación de postgrado tanto en el área de sistemas de información como en el área biomédica, así como también experiencia en el ámbito profesional en lo que refiere al desarrollo de sistemas.

Se incorporan tres alumnos becarios de la carrera Licenciatura en Sistema de la Facultad de Ciencia y Tecnología de la UADER

Referencias

[1] M.G. Piattini, J.A. Calvo-Manzano, J. Cervera, L.Fernandez. Análisis y Diseño de Aplicaciones informáticas de Gestión. Una perspectiva de Ingeniería del Software. 2007, ISBN 978-84-7897-776-5. RA-MA Editorial

[2] M. E. Llorente, Aldo Daniel Sigura, Javier Besso, Alejandro Hadad, Bartolomé Drozdowicz. “Proceso de Diseño basado en Casos de Uso para un Datawarehouse Clínico”. CACIC 2012

[3] M. E. Llorente, Aldo Daniel Sigura, Javier Besso, Ernesto Mangia, Alejandro Hadad, Maria Manzini, Noralí Quijada y Bartolomé Drozdowicz. “Propuesta de un Modelo Multidimensional para Datawarehouse sobre pacientes diabéticos” WICCC 2014, Ushuaia, Tierra del Fuego.

[4] M. E. Llorente, Aldo Daniel Sigura, Alejandro Hadad, Bartolomé Drozdowicz. “Sistema de soporte a la toma de decisiones basado en datawarehouse para pacientes

diabéticos.” WICC 2012

[5]. M. E. Llorente, A. Sigura, J. Besso, E. Mangia, A. J. Hadad, M. Mancini, N. Quijada, B. Drozdowicz. “Aplicación de prototipo para el Modelo Multidimensional de un Datawarehouse sobre pacientes diabéticos”. WICC 2015, Ushuaia, Tierra del Fuego.

[6] E.Mangia, O.Denis, M.E.Llorente, J.Besso, A.Sigura, B. Drozdowicz, A.Hadad. “Proceso iterativo de diseño del modelo multidimensional de soporte a datos clínicos de pacientes diabéticos”. Ventana Informática. Universidad de Manizales.2015.ISSN 0123-9678