

Desarrollo de Ontologías para capturar el conocimiento experto en Modelado del Sistema Endocrino de Pacientes Diabéticos

Juan José Bora¹, Maximiliano Benegui¹, Pamela Viale^{1,2} y Marta Basualdo^{1,2,3}

¹Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Av. Pellegrini 250, S2000BTP Rosario, Argentina.

²Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas
CONICET - UNR - AMU

Ocampo y Esmeralda, S2000EYP Rosario, Argentina

³GIAIP-FRRO-UTN

Zeballos 1341 S2000BTP Rosario, Argentina

borajuanjo@gmail.com

maxibenegui@yahoo.com.ar

viale@cefasis-conicet.gov.ar

basualdo@cefasis-conicet.gov.ar

Abstract. El número de pacientes con Diabetes Mellitus (DM) ha crecido a nivel mundial de manera exponencial a lo largo de los años. Esto justifica el enorme esfuerzo que la comunidad internacional viene realizando para abordar multidisciplinariamente esta enfermedad mediante el desarrollo de tecnologías adecuadas. En este trabajo se presenta una extensión de la ontología denominada OntoCAPE para capturar el conocimiento del modelado de tipo compartimental del sistema endocrino de pacientes diabéticos. El objetivo es además tener una versión preliminar de una plataforma que brinde servicios multidisciplinarios de tipo *e-health*. Expertos de diferentes áreas podrán verter sus conocimientos en medicina, nutrición e ingeniería. Se presenta un detalle de la nueva conceptualización introducida a OntoCAPE, inicialmente desarrollada para su uso en Ingeniería Química y ahora se la extiende para su uso en sistemas biológicos. A su vez se incorporó la integración con el programa MATLAB a fin de poder generar el código en el marco de su herramienta *s-function* y poder testear información de pacientes específicos en modo dinámico.

Keywords: Diabetes, Sistema Endocrino, ingeniería del conocimiento, ontologías, modelado, OntoCAPE, web ontology language.

1 Introducción. Diabetes Mellitus.

La Diabetes Mellitus es una enfermedad crónica, que en los últimos años se ha convertido en una epidemia. La enfermedad aparece cuando el páncreas no puede producir suficiente insulina (Tipo I) o cuando la misma no presenta las propiedades

necesarias para que resulte eficaz para el organismo (Tipo II). La hormona insulina es la que permite que la glucosa, proveniente de los alimentos ingeridos, se metabolice e ingrese a las células del cuerpo y pueda convertirse en energía necesaria para los músculos y tejidos.

El efecto de la Diabetes no controlada se manifiesta en casos de hiperglucemia y hipoglucemia. La hiperglucemia (aumento de la glucosa en sangre), con el tiempo daña gravemente muchos órganos y sistemas, especialmente el nervioso y los vasos sanguíneos. Por otro lado, la hipoglucemia (disminución de glucosa en la sangre) puede producir la muerte del individuo y por lo tanto requiere mayor atención para evitar episodios de estas características.

Según datos reportados por organismos internacionales como la Organización Mundial de la Salud (OMS), en la actualidad el 9% de los adultos mayores de 18 años padece diabetes, y sólo en 2012 la enfermedad causó 1.5 millones de muertes. Se prevé que para el año 2030 la diabetes será la séptima causa de muerte [1].

El gasto sanitario mundial para tratar la diabetes y prevenir complicaciones alcanzó al menos 548.000 millones de USD en 2013. Este número representó el 11% del gasto sanitario mundial de ese año. Para 2037, se prevé que este número supere los 592.000 millones de USD [2].

La situación en Argentina, también manifiesta datos alarmantes. En 2010 el Dr. Juan José Gagliardino, miembro del CONICET, señaló que “La prevalencia de esta enfermedad en adultos mayores en nuestro país es del 8,5%; es decir que más de 2.000.000 de argentinos la padecen” [3]. Además, remarcó que el tratamiento de un diabético complicado cuesta de 2 a 4 veces más que uno que no lo está. Gagliardino destacó además las ventajas del programa que llevó a cabo, llamado *Program for the Prevention, Care and Treatment of People with Diabetes (PROPAT)*. Las ventajas obtenidas del mismo fueron: la disminución en la tasa y en la duración de la hospitalización en personas que sufren de diabetes comparada con las que no lo sufren. Esto trae ventajas económicas ya que la hospitalización representa aproximadamente el 50% del costo total de la atención. Los pacientes con un pobre control glucémico tienen más del doble de internaciones durante un período de tres años, en comparación con aquellos que tienen un buen control de la glucemia.

El tratamiento de la enfermedad Diabetes Mellitus es complejo, riesgoso y demandante en exceso. Para mejorar la calidad de vida del paciente a largo plazo, es necesario que el mismo cuente con la mayor cantidad de información posible acerca de la enfermedad que padece. Al mismo tiempo, es importante también que dicha información se encuentre al alcance de todo el equipo de profesionales que se encuentra involucrado en el tratamiento, como ser: médicos, nutricionistas, diabetólogos, técnicos o personal ingenieril. Por otra parte, es importante también que el paciente pueda reducir el número de intervenciones médicas presenciales, lo cual disminuirá notablemente los costos del tratamiento.

Resulta, entonces, de especial importancia el desarrollo de herramientas computacionales y nuevas bases de conocimiento que podrían integrarse a programas como el PROPAT, con la ventaja de facilitar notablemente el manejo de la información de pacientes diabéticos, mejorar el nivel educativo de los mismos contribuyendo a construir su propio modelo de comportamiento para fortalecer actitudes de propio auto-

control, disminuir la dependencia de los expertos involucrados y llevar adelante una buena parte del programa con una fuerte asistencia de personal técnico capacitado. Así mismo, con el desarrollo de las mencionadas herramientas computacionales del conocimiento, contribuiría al mejor seguimiento del paciente y sería esperable la disminución del número de intervenciones médicas presenciales.

2 Modelo Matemático del Sistema Endocrino Humano. Simulador UVA/Padova [4] [5].

El modelo desarrollado y patentado por las Universidades de Virginia y Padova, llamado *Simulador UVA/Padova*, se presenta en la Fig. 1, describe la relación entre la glucosa en plasma y la concentración de insulina, y flujos de glucosa e insulina en el sistema endocrino. El modelo matemático es factible de ser utilizado para emular el sistema endocrino de un humano saludable, prediabéticos, diabéticos Tipo II y diabéticos Tipo I. Debido a esto es uno de los pocos modelos que fue validado con datos clínicos experimentales, la versión del paciente diabético tipo I fue aprobada por la Administración de Alimentos y Drogas de Estados Unidos, *Food and Drugs Administration* (FDA), como sustituto de las pruebas en animales para testeos pre-clínicos de algoritmos de control.

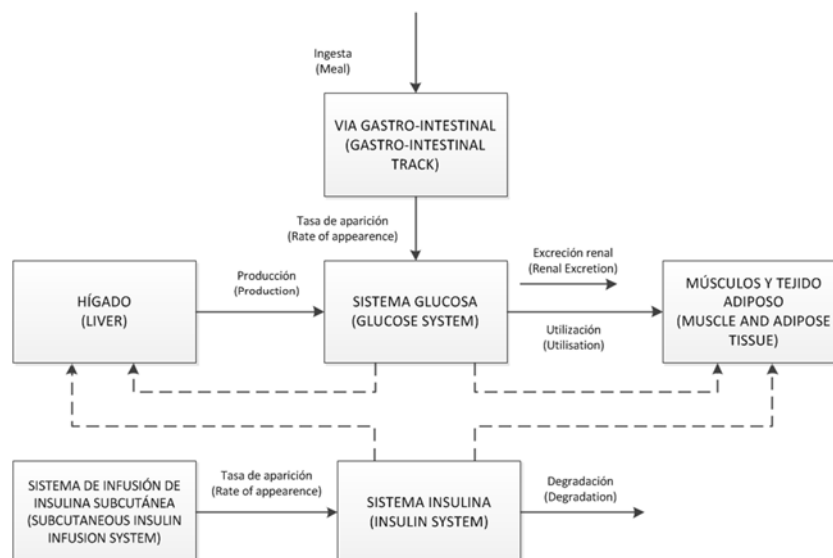


Fig. 1. Esquema del sistema glucosa-insulina del Simulador UVA/Padova. Las líneas continuas representan flujo de masa, y las líneas a trazos representan flujo de señales [5].

Vale mencionar que el modelo completo cuenta con una base de datos de 300 sujetos diabéticos tipo I (100 adultos, 100 adolescentes y 100 niños) pero sólo está disponible para los miembros de la Juvenile Diabetes Research Foundation (JDRF) Artifi-

cal Pancreas Consortium. Este modelo permitió simular el efecto dinámico de la glucosa exógena y de la dosis de insulina bajo diferentes pruebas. En la Fig. 1 se presenta el modelo correspondiente a un paciente diabético, donde se tiene en cuenta la infusión de insulina sub-cutánea.

El modelo matemático de tipo compartimental consta de un sistema de ecuaciones diferenciales, las cuales describen la dinámica de los Sub-Sistemas de Ingesta, Glucosa, Insulina y Espacio Subcutáneo del Sistema Endocrino Humano.

El Simulador UVA/Padova ha sido implementado íntegramente en la plataforma MatLab/Simulink [6], mediante la utilización de una S-FUNCTION [6] de MatLab.

3 OntoCAPE [7]

OntoCAPE es una ontología de gran alcance para el dominio de la ingeniería de procesos asistida por computadora. Su nombre tiene el siguiente significado: ‘Onto’, hace referencia a que se trata de una ontología y ‘CAPE’ es la sigla en inglés de *Computer-Aided Process Engineering*¹. Ha sido desarrollada en la *RWTH Aachen University* en Alemania [7]. OntoCAPE, a su vez, es accesible bajo la licencia GNU y tanto su representación formal como informal puedan ser libremente descargadas desde la Web [8]. Por último, cabe aclarar que OntoCAPE se encuentra implementada en lenguaje de ontologías OWL (*Web Ontology Language*) [9], y se optó por conservar esta concepción si bien a futuro se podría evaluar otro lenguaje, como OWL2, por ejemplo.

OntoCAPE está particionada en 62 sub-ontologías, que pueden ser usadas individualmente o como un producto completo integrado en conjunto. Estas sub-ontologías se organizan en diferentes capas de abstracción, que separan conocimiento general de conocimiento sobre dominios y aplicaciones en particular. Las capas superiores tienen la característica de una ontología superior, cubriendo temas generales, como por ejemplo teoría de sistemas, cantidades o unidades. Las capas inferiores conceptualizan el dominio de ingeniería de procesos químicos, cubriendo temas específicos de dominio como por ejemplo materiales, reacciones químicas o unidades de operación.

OntoCAPE es una ontología pensada para la representación de conocimiento de procesos químicos y de la ingeniería de procesos en general. Resulta de gran utilidad que sus diseñadores hayan tenido en mente la posibilidad de futuras adaptaciones y extensiones de la misma, proponiendo un diseño abierto y flexible. Esto ha imaginado la extensión de OntoCAPE para el modelado de sistemas biológicos, ya que se comportan de manera similar o pueden ser modelados como reacciones químicas. Su representación física, a su vez, puede asimilarse a equipamientos de una planta química.

Si bien OntoCAPE es una ontología desarrollada para la industria química, tiene la ventaja de presentar un conjunto de herramientas conceptuales que permiten modelar el Sistema Endocrino. La representación del mismo junto con la modelización del perfil de ingesta del paciente proveerá información de vital importancia para la multi-plataforma *e-health* que se tiene como objetivo a largo plazo. Esta plataforma podrá

¹ Ingeniería de Procesos Asistida por Computadora.

sustentarse en dichos modelos para inferir valiosa información para la toma de decisiones tendientes a mejorar el tratamiento, y la calidad de vida de un paciente con DM.

Podría considerarse a OntoCAPE como una base de conocimientos la cual facilita la carga de datos, parámetros y resultados, asociados entre sí mediante un marco semántico. Esto favorece el entender y el compartir dicha información asegurando que la información llegue de manera adecuada a cada persona o software involucrado. Todos los datos cargados en una ontología y toda la información semántica volcada en ella, ya sea especificada o inferida, puede ser accedida mediante programas externos, desarrollados por expertos en el área de la ingeniería de conocimiento.

4 Representación del Sistema Endocrino Humano en OntoCAPE. Aspectos descriptivos del trabajo realizado [10].

Se entiende que el Sistema Endocrino Humano es un Sistema de Procesos Químicos Biológico, lo cual es considerado un caso particular de un Sistema de Procesos Químicos. Es por ello que se ha escogido OntoCAPE como ontología raíz, para hacer re-utilización de sus definiciones y reglas para encajar el modelado del Sistema Endocrino Humano.

Para la representación del modelado del Sistema Endocrino humano en términos de ontologías, se utiliza el módulo `chemical_process_system` de OntoCAPE, accesible como una sub-ontología. Dicho módulo presenta la capacidad de realizar el modelado mereotopológico completo del Sistema Endocrino Humano, en términos del correspondiente al Simulador UVA/Padova. Es decir: con la creación de clases, instancias y propiedades capaces de definir el modelado compartimental correspondiente al simulador UVA/Padova y asociar al mismo semánticamente con su respectivo sistema de ecuaciones diferenciales no-lineales.

El conocimiento volcado en la ontología puede ser atomizado, tanto cuanto se desee. Para el presente caso no solamente se modela la concepción mereotopológica, sino que también se detallan flujos de masas, de señales y el concepto de sub-sistemas y super-sistemas.

Respecto del sistema de ecuaciones diferenciales, cabe la aclaración de que las ecuaciones correspondientes son convenientemente representadas mediante la estructura de datos *binary tree*. Dicha estructura de datos permite asociar un *individual* a una parte izquierda (*left child*) y una parte derecha (*right child*) de una ecuación, unidas por un operando (*root node*). La parte derecha o izquierda puede ser, a su vez, o bien un operando definitivo de la ecuación (denominado ‘hoja’, en inglés *leaf*), o bien una ecuación subsiguiente (*internal node*). Con la utilización de un programa implementado para tal fin, se extrae la información de las ecuaciones de la ontología, y se las traduce a lenguaje MatLab. Esto permite ser utilizadas en una estructura de diagrama de bloques de este programa, que asocia un sistema a un bloque mediante la utilización de un módulo denominado S-FUNCTION. Por motivos que excede al alcance del presente trabajo, no se aborda aquí dato alguno acerca del programa im-

plementado para la lectura e interpretación de los *binary tree*, y tampoco se dirá nada al respecto de la plataforma de simulación implementada como diagrama en bloques de MatLab/Simulink.

Cabe aclarar que el modelo presenta una parte destinado a las ingestas de glucosa y permite anuncios de éstas. Dichos conceptos han sido preliminarmente introducidos en la ontología con miras a integrarse a futuro con ontologías dedicadas específicamente a nutrición. Se prevé además que la ontología pueda ser extendida, para que la misma forme parte de una ontología de mayor escala, capaz que ser el nexo de información compartida entre profesionales de la medicina y de nutrición. El campo de extensión de la ontología puede, a su vez, ser sucesivamente extendido a más áreas profesionales.

A los efectos de abordar una descripción más profunda, se aclara que todos los procesos del Sistema Endocrino Humano se describen en la ontología con *individuos* que representan tanto su descripción física como funcional. Finalmente, la descripción funcional de cada proceso se asocia con el correspondiente sub-sistema de ecuaciones diferenciales, el cual describe matemáticamente su funcionamiento. Dicho de otra forma, para la implementación del modelado en términos de ontologías, se crearon estructuras de clases, *individuos* y propiedades, a los efectos de contar con una descripción funcional y física de todos los procesos del Sistema Endocrino Humano.

Por ejemplo, el individual **stomach**, perteneciente a la clase **Stomach**, la cual es, a su vez, subclase de **Reactor** (la cual es una clase de la estructura básica de OntoCAPE), se ha establecido la propiedad en la ontología, que realiza (**realizes**), el proceso **stomach-process**, lo cual es una instancia de la clase **ProcessStep** (escalón de un proceso), que representa un sub-proceso de un proceso completo global. Este **stomach-process**, a su vez, está relacionado, en la ontología, con su respectivo sub-sistema de ecuaciones diferenciales.

Para lograr el nivel de detalle descrito anteriormente, se definieron *clases*, *individuos* y *propiedades de objeto* en la ontología OntoCAPE base, lo cual se considera haber conformado una extensión de la misma. La cantidad de elementos de ontología definidos para realizar dicha extensión, asciende a más de 600, y por razones de espacio del presente trabajo no podrán describirse en su totalidad.

Principalmente, todo el modelado del Sistema Endocrino Humano se realizó teniendo en cuenta las reglas de modelado de OntoCAPE, en el módulo **chemical_process_system**. Dichas reglas se describen en las *Fig.2* y *Fig.3*. Se tiene en cuenta que para todo tipo de representación de Sistemas de Procesos Químicos (**ChemicalProcessSystem**), se considera que éstos están compuestos por un determinado número de Unidades del Proceso (**ProcessUnit**). Así, se describe el Sistema Endocrino Humano como un *chemical process system*, y se hace hincapié en representar a los Sistemas Ingesta, Glucosa, Insulina, Musculación e Hígado como sub-sistemas del Sistema Endocrino, todo esto con una profunda capa de abstracción semántica. Más aún, todos los parámetros relacionados con la operación dinámica de los distintos procesos, tienen una descripción en la ontología, y están asociados a su valor numérico y unidad de medida, de acuerdo a las especificaciones del Simulador.

En la ontología se definen también conceptos para representar flujos de señales y flujos de masas, de una manera discriminada.

Para los trabajos de edición de ontologías, se utilizó *Protégé 4.3 (build 304)* [11]. Para efectos de testeo de consistencia del conocimiento modelado en la ontología se utilizó el razonador *HermiT 1.3.8* [12]; el cual ha sido el razonador de ontologías que mejor rendimiento y resultados ha mostrado en los testeos. Con la utilización del razonador se accede a conocimiento complementario modelado en la ontología, representado mediante inferencias.

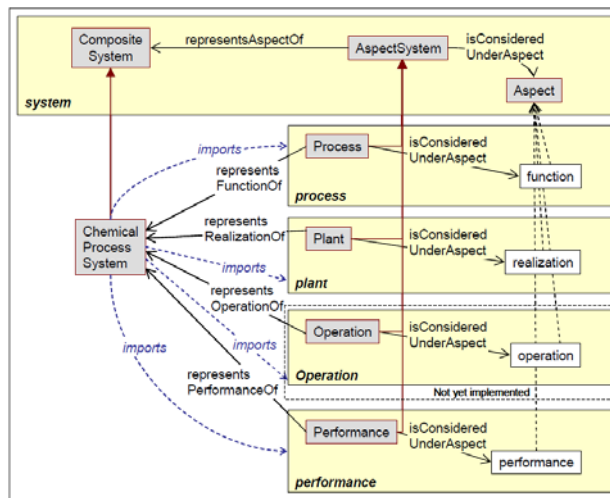


Fig. 2. Presentación de la consideración de *aspect system* para la parte de procesamiento de un *chemical process system*.

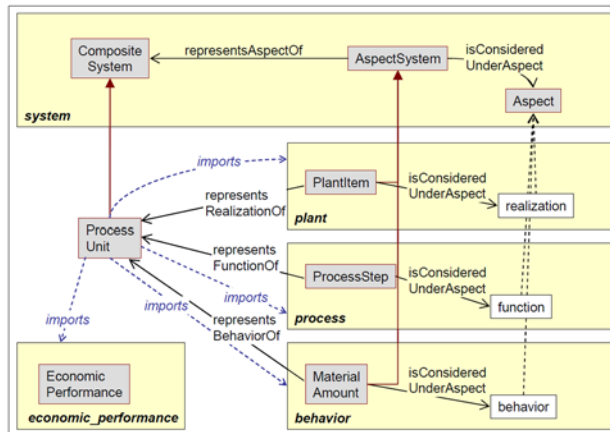


Fig. 3. Presentación de consideraciones de *aspect system* para *process unit*.

La metodología de trabajo adoptada para llevar adelante la tarea de modelado fue, en primer lugar, el estudiar y generar conocimiento acerca del módulo **chemical_process_system** de OntoCAPE. Posteriormente, realizar un relevamiento acerca de qué módulos de **chemical_process_system** son aptos para el modelado que se está intentando hacer. Depuración de la ontología, eliminando el exceso de información, y así, aquellos módulos que no son necesarios para la descripción del sistema que se desea modelar. Extensión de las clases **ProcessStep**, **PlantItem**, **ProcessUnit**, **AspectSystem**, **Plant**, **Process**, **ChemicalProcessSystem**, etc, a fin de lograr la estructura de clases deseada y apta para la descripción del proceso específico del modelo matemático implementado en el Simulador UVA/Padova. Posteriormente, se realizó la creación de todas las instancias de las clases y la asociación entre ellas mediante propiedades de objeto de la ontología. La mayoría de las propiedades de objeto utilizadas, son parte del paquete OntoCAPE por defecto, sin embargo, un gran número de propiedades de objeto han debido crearse.

5 Implementación [10]

Para la implementación de la ontología se escribió un programa en *Java*, capaz de obtener, de la ontología, el conocimiento del sistema de ecuaciones diferenciales del Simulador UVA/Padova. Esta tarea se realiza con métodos que analizan la estructura de datos de *binary tree*. Una vez recuperadas las ecuaciones, y haciendo utilización de una serie de reglas de modelado, el programa construye la S-FUNCTION del Simulador UVA/Padova, a los efectos de crear el ambiente adecuado para una simulación dinámica de los pacientes de la base de datos del Simulador, ahora cargados en la ontología.

Con dicho programa se crea también el escenario para la simulación, en *MatLab/Simulink*, del mencionado sistema de ecuaciones, y se muestra en pantalla la evolución de glucosa en sangre de un paciente determinado, de la base de datos del 30 pacientes disponible. Todos los parámetros del modelo matemático relacionados con la característica particular del Sistema Endocrino de cada paciente son también cargados en la ontología, y recuperados con el programa, para la carga de los mismos en el espacio de trabajo de la herramienta de cálculo. Esto significa una contribución de gran utilidad respecto del punto de partida que representa la estructura de OntoCAPE, dado que la misma provee escasa información para la implementación de simulaciones dinámicas de los procesos descriptos.

6 Conclusiones

Se ha puesto de manifiesto, en este trabajo, que es posible incorporar la rama semántica a software dedicado al modelado del sistema endocrino de pacientes diabéticos. El conocimiento modelado en la ontología puede ser llevado a niveles de aún

mayor abstracción, dado que OntoCAPE permite la (re)utilización de sí misma, y así también la ontología resultante con el modelado del Sistema Endocrino Humano.

Con la utilización del razonador informático se puede corroborar en todo momento la consistencia de la ontología, a modo de no incurrir en errores semánticos. Pueden también agregarse nuevas *propiedades de objeto* y *clases*, para profundizar en el nivel de detalle de los procesos modelados.

El conocimiento acerca del funcionamiento del Simulador UVA/Padova, representado en la ontología mediante la descripción en *binary tree* del sistema de ecuaciones diferenciales, ha podido ser leído, y utilizado adecuadamente para la implementación de simulaciones *in-silico* de la dinámica de la glucosa en sangre de pacientes determinados. Gracias a ello, con el software implementado, se captura en la ontología todo el conocimiento acerca del Simulador UVA/Padova, y se puede, desde la misma ontología, efectuar simulaciones dinámicas en herramienta de cálculo. No se presentan en la actualidad, casos similares de extensión de OntoCAPE para el manejo de simulaciones dinámicas de pacientes diabéticos en herramientas de cálculo, como lo presentado aquí.

6.1 Trabajos Futuros.

Como se dijo anteriormente, nada relacionado con la ingesta oral de glucosa ha sido modelado en detalle en la presente versión de la ontología. Representa un gran desafío a futuro el sintetizar una ontología que permita el cálculo de una dosis oral de glucosa ingerida, mediante la especificación de alimentos comerciales que el paciente incorpora en su organismo.

Se espera extender la versión ampliada de OntoCAPE para el tratamiento de sistemas biológicos con miras a ser utilizada como núcleo de una plataforma de lenguaje compartido en telemedicina para profesionales de diferentes áreas abocadas al tratamiento de pacientes diabéticos.

7 Referencias

1. Organización Mundial de la Salud, Diabetes: Factsheet, <http://www.who.int/mediacentre/factsheets/fs312/en/>.
2. International Diabetes Federation: IDF Diabetes Atlas 6th Edition. Brussels, Belgium (2013).
3. Entrevista al Dr. Juan José Gagliardino: "La prevalencia de la diabetes en la población argentina es de 8.5%", <http://www.unirse.com.ar/edic%20165.htm>.
4. Campetelli, G.: Desarrollo de Bio-modelos Computacionales para Asistir en la Toma de Decisiones Tendientes a Mejorar la Calidad de Vida de Pacientes Diabéticos. CIFASIS-CONICET; Facultad de Ciencias Exactas, Ingeniería y Agrimensura; Universidad Nacional de Rosario, Rosario, Argentina (2014).
5. Colmegna, P.: Simulation & Control in Type I Diabetes. Instituto Técnico de Buenos Aires, Buenos Aires, Argentina (2014).

6. MatLab, Simulink y MatLab S-FUNCTION Websites,
<http://www.mathworks.com/products/matlab/>,
<http://www.mathworks.com/products/simulink/> y
<http://www.mathworks.com/help/simulink/sfg/what-is-an-s-function.html>.
7. Marquardt, W., Morbach, J., Wiesner, A., Aidong, Y.: OntoCAPE A Re-Usable Ontology for Chemical Process Engineering. Berlin: Springer RWTHeedition (2010).
8. The user community of the re-usable Ontology for Chemical Process Engineering,
<http://www.ontocape.org/>
9. Web Ontology Language Overview, <http://www.w3.org/TR/owl-features/>.
10. Bora, J., Benegui, M.: Utilización de Ontologías para Capturar el Conocimiento Experto en Modelado del Sistema Endocrino Humano y su Control (Proyecto de Ingeniería). Facultad de Ciencias Exactas, Ingeniería y Agrimensura, en colaboración con CIFASIS-CONICET. Rosario, Argentina (2015).
11. Protégé Website, <http://protege.stanford.edu/>.
12. Hermit OWL Reasoner Website, <http://hermit-reasoner.com/>.