

# Capítulo 2: Estado del arte de los sistemas de gestión del conocimiento en entornos sanitarios

## 2.1 Introducción

Los avances en las TIC (Tecnologías de la Información y de las Comunicaciones) han propiciado cambios significativos dentro de los sistemas sanitarios actuales. Entre otras aportaciones, este desarrollo ha posibilitado que un gran número de países hayan implantado [8] sistemas EHR (Registro Electrónico Sanitario), permitiendo el uso de ontologías que intentan alcanzar un determinado nivel de interoperatividad semántica entre diferentes instituciones sanitarias.

Sin embargo, a pesar de que se han desarrollado infraestructuras tecnológicas para compartir recursos de todo tipo, la mayoría de usuarios que hacen uso de ellas se enfrentan a una cantidad de información demasiado cuantiosa [9], y en constante crecimiento. Prueba de esto último, es la explosión de información surgida en los últimos años a partir de la revolución genética, que ha favorecido la aparición de nuevos modelos de datos cada vez más complejos. La interacción entre la bioinformática y este tipo de innovaciones científicas incrementarán en un futuro el interés por una eficiente gestión del conocimiento para adquirir, entender e interpretar de la forma más eficiente posible, toda la ingente cantidad de información clínica disponible, con el objetivo de obtener el máximo conocimiento posible [10].

A pesar de todo, la importancia del conocimiento clínico dentro del entorno sanitario se ha convertido ya en una realidad, y prueba de ello es el aumento exponencial de la literatura asociada a este tema en los últimos años. Por ejemplo, la MEDLINE (Análisis de la literatura médica y Sistemas de Recuperación de Datos) [11] contaba en 2002 con 11.7 millones de citas [12], mientras que actualmente esa cifra ha sido ampliamente rebasada y ronda ya los 17 millones [13].

Antes de definir en primer lugar el concepto de sistemas de gestión del conocimiento, es necesario distinguir los conceptos de conocimiento y de gestión del conocimiento en términos generales. El conocimiento como tal, es una notación abstracta que lleva utilizándose desde hace siglos, y puede definirse como “cualquier opinión personal justificada que incrementa la capacidad de los individuos para tomar acciones efectivas” [14]. En la literatura existente dentro de los sistemas de información, la definición de este término suele estar acompañada también por una distinción entre conocimiento, información y datos.

En este sentido, Vance [15] define la información como datos que son interpretados dentro de un contexto con significado, y el conocimiento es la información que ha sido comprobada como cierta. Por otra parte, Maglitta [16] sugiere que los datos contienen un conjunto de hechos y números, la información son datos procesados, y el conocimiento es información que puede ser utilizada para alcanzar un determinado fin. La clave para distinguir correctamente entre información y conocimiento no se encuentra en el contenido, estructura, exactitud o utilidad de los mismos. Más bien se podría decir que el conocimiento es información interpretada por un individuo en cuestión, y por tanto se trata de una visión subjetiva y personalizada en referencia a unos hechos, procedimientos, conceptos, interpretaciones, ideas, observaciones y juicios [17].

El conocimiento puede ser clasificado en dos tipos: explícito y tácito. El conocimiento explícito reúne los hechos, reglas, relaciones y políticas codificadas a través de un lenguaje formal, y que son compartidas sin necesidad de discusión [18]. Este tipo de conocimiento puede estar

contenido en manuales, informes o bases de datos electrónicas [19]. En contraste con esto, el conocimiento tácito se basa en ideas personales o intuiciones, es decir, está relacionado con la experiencia personal que aporta un individuo, por lo que dicho conocimiento presenta un carácter subjetivo y requiere de una interpretación concreta para entender su significado [10]. Aunque ambos tipos de conocimientos son importantes, el conocimiento tácito es más complicado de identificar y manejar, y más aún dentro del ámbito sanitario [20].

La categorización de-facto para clasificar conocimiento en tácito o explícito está basada en dos tipos de estrategias: de codificación, y de personalización [21]. La estrategia de personalización enfatiza en el conocimiento tácito que reside en una organización, mientras que la estrategia de codificación está enfocada en el uso de una infraestructura tecnológica que provea soporte para permitir la identificación, almacenamiento y transferencia del conocimiento disponible [10].

En consecuencia, la gestión del conocimiento se refiere al proceso sistemático y organizado para adquirir, organizar y comunicar conocimiento explícito y tácito entre los componentes heterogéneos de una determinada organización, de tal forma que su uso permita una mayor efectividad y productividad en el trabajo. Los sistemas de gestión del conocimiento son los encargados de realizar esta labor.

En el ámbito sanitario, la gestión del conocimiento va a venir dada por la producción, distribución, almacenamiento, evaluación, transferencia y puesta en práctica del conocimiento clínico disponible [21], con el objetivo de mejorar la calidad asistencial al paciente, la provisión de servicios, la gestión de los recursos disponibles, y como meta final, incrementar el nivel de salud de los ciudadanos [22].

Esta gestión del conocimiento está asociada con la información que se le ofrece al mismo, el conocimiento que se pueda aportar de manera individual o colectiva a organizaciones sanitarias, y la influencia de las nuevas tecnologías en aspectos como la relación médico-paciente o la organización de la asistencia sanitaria [23].

Más concretamente, en [22] identifican un paradigma de gestión del conocimiento en entornos sanitarios que reúne un conjunto de procesos para transferir y explotar conocimiento de fuentes de datos internas y externas. Son 4 procesos interrelacionados: (1) creación o adquisición de conocimiento, (2) integración o conversión de conocimiento, (3) aprovechamiento o aplicación de conocimiento y (4) protección de conocimiento. El nuevo conocimiento es creado o adquirido a partir de la interacción con el paciente, convertido para facilitar el acceso y uso al mismo, y finalmente aprovechado para aplicarlo en otro tipo de escenarios relevantes. Esta serie de procesos de gestión del conocimiento se ve ilustrado en la Figura 2.1.



Figura 2.1. Procesos de gestión del conocimiento en entornos sanitarios

Desde este punto de vista, se han presentado otros paradigmas para la gestión del conocimiento en entornos sanitarios, pero en la práctica, ninguno ha tenido éxito [24], debido a la inhabilidad para combinar procesos organizativos sanitarios con tecnología, entre otras cuestiones [25].

A pesar de todo, en los últimos años se ha incrementado el interés en investigar la naturaleza y utilización del conocimiento clínico por medio de teorías de gestión del conocimiento, metodologías y entornos de trabajo, teniendo en cuenta el principal objetivo, que es mejorar la calidad de cuidado del paciente. Sin embargo, este propósito se torna muy complejo de alcanzar, ya que la gestión del conocimiento necesita soporte continuo debido a la evolución del paciente y a sus correspondientes procesos de cuidado. A esto hay que añadir, que los sistemas de gestión del conocimiento clínico tienen que hacer frente a diferentes heterogeneidades en cuanto a recursos de conocimiento disponible, disparidad entre las instituciones, organismos, y demás organizaciones implicadas en la diseminación de este conocimiento, con diferentes capacidades, terminologías, necesidades y expectativas, y otros aspectos relacionados [8].

Por último, señalar que recientemente, y gracias al auge de las TIC, se ha introducido un nuevo concepto denominado UbiCare (Ubicuidad Sanitaria) [26], que presenta una nueva perspectiva a la hora de acceder a los servicios sanitarios, proporcionando una mayor personalización y una mejora en la atención al paciente y en su tratamiento. En este sentido, en [26] especulan sobre el posible impacto de este nuevo paradigma dentro de los sistemas de gestión del conocimiento sanitario actual.

Dentro de los sistemas de información sanitarios, existen diferentes áreas en donde la gestión del conocimiento ha adquirido una fuerte implicación, y por tanto, se ha optado por un análisis breve de cada uno de estos campos, para alcanzar una abstracción que permita ofrecer una visión más completa y trate de simplificar la complejidad del tema.

## 2.2 Registro electrónico del paciente

El concepto de EHR (Registro Electrónico Sanitario) está íntimamente relacionado al de los sistemas de información dentro de los entornos sanitarios, y por tanto, a la de la gestión del conocimiento. Estos registros contienen el conjunto completo de información de un paciente en formato electrónico. El propósito principal del EHR es documentar, recuperar, transmitir, enlazar y procesar información para la entrega de conocimiento y soporte a la decisión clínica que mejore la eficiencia y seguridad de los servicios sanitarios. Existen otros usos secundarios del uso del EHR, como los relacionados con educación, investigación, gestión de calidad o estadísticas.

Las características más importantes del EHR son las siguientes [27]:

- Se encuentra centrado únicamente en el paciente, e incluye toda la información relevante sobre el mismo. Por tanto, difiere de otro tipo de contenedores de información centrado en un centro sanitario en particular, que contiene registros episódicos de los pacientes para ese centro.
- Incluye información no solo del pasado o presente del paciente, sino que también presenta información prospectiva del estado de salud, es decir, el EHR contiene observaciones (lo

que ha ocurrido), opiniones (decisiones sobre lo que está ocurriendo), y planes de cuidado (lo que debería ocurrir en el futuro).

- El nivel de abstracción del EHR es generalista, es decir, recursos específicos tales como imágenes médicas o algoritmos de soporte para la decisión clínica no forman parte del EHR, sino que existen otro tipo de estándares para gestionar este tipo de información.
- El EHR es una fuente de información que puede ser útil no solo para el personal sanitario y los sistemas de soporte a decisiones clínicas, sino también es empleado con propósitos de investigación, gubernamentales, estadísticos, y en otro tipo de contextos.

Sin embargo, mientras los HIS (Sistemas de Información Hospitalarios) o los CIS (Sistemas de Información Clínicos) no utilicen estándares consensuados que implementen el EHR, y que faciliten el intercambio electrónico de los datos, no es posible que la información esté disponible en el punto de atención donde se encuentra el paciente, independientemente de la institución prestadora de servicios de salud donde sea atendido.

Dentro de los estándares informáticos de salud más desarrollados que implementan el EHR, el más extendido es el HL7 (Health Level Seven) [28], que cuenta con especificaciones de mensajes, documentos electrónicos y vocabularios controlados para dominios de salud tales como: CDA (Arquitectura Clínica de Documentos), registros médicos, laboratorio, medicación, diagnóstico e integración DICOM (Estándar de Imagen Digital y Comunicaciones en Medicina), prestación de atención, entre otras cuestiones.

Otro de las fundaciones que ha tomado gran relevancia en los últimos años en el desarrollo del EHR, ha sido el openEHR [29], que se trata de una especificación que describe cómo debe gestionarse, almacenarse e intercambiarse la información sanitaria contenida en el EHR. La aportación más interesante del openEHR es la de dar respuesta a la necesidad de interoperatividad semántica, esto es, la capacidad mutua de los computadores y del software para entender el significado de la información. Para ello, la arquitectura de información del openEHR consta de dos niveles de organización que intenta alcanzar esta abstracción. Por un lado, define un modelo de referencia que describe las características de las entradas de información de los registros de los pacientes, cómo se agregan y relacionan entre ellas, y la información de contexto requerida para satisfacer los requerimientos éticos, legales y de autoría de los datos clínicos. Por otra parte, un modelo de arquetipos, que limita las clases definidas en el modelo de referencia para dominios clínicos particulares, organizaciones, y contextos operacionales.

Por último, cabe destacar la norma 13606, constituida en el seno del TC251 del CEN (Comité Europeo de Normalización), y que toma prestada algunas de las características del openEHR. Básicamente define una arquitectura de información clínica y su comunicación entre sistemas de información heterogéneos. De la misma manera que el openEHR, la norma EN13606 también hace uso de un modelo de referencia para la representación de la información, y un modelo de arquetipos para el conocimiento.

## 2.3 Lenguaje y codificación de la información

La representación de los datos médicos es uno de los fundamentos en los que se sustenta el campo de la informática médica. La información recogida de fuentes de datos heterogéneas, como por ejemplo, de analíticas, dispositivos portables que monitorizan constantes vitales del paciente, resultados de laboratorio etc., son los que conducen a la creación de nuevo conocimiento. Los sistemas de codificación de información se encargan de extraer los datos relevantes de todas estas

fuentes heterogéneas, basándose en modelos, métodos, y técnicas existentes en el dominio médico.

Una de las técnicas que se utilizan para obtener una información de mayor calidad, y por tanto, para la extracción de conocimiento más útil, es la integración de los recursos de información ya existentes en una institución, con otros pertenecientes a fuentes externas.

Por ejemplo, particularizando para el entorno biomédico, la ontología GO (Ontología Gene) [30] ha constituido una fuente de información muy valiosa para investigadores que trabajan en temas relacionados con el genoma humano [32]. Otro ejemplo es la ontología MeSH (Encabezamiento de Temas Médicos) [31], que ha sido aplicada con éxito para expandir términos de las consultas en sistemas de recuperación de información [32].

Dentro del conocimiento biomédico, el UMLS (Sistema de Lenguaje Unificado de Modelado) [33] constituye un repositorio o metatesauro con una extenso vocabulario de conceptos relacionados con la biomedicina, así como un cuantioso número de relaciones entre dichos conceptos.

UMLS está formado por tres componentes básicos:

- Una fuente de conocimiento que posee conceptos del dominio de la biomedicina y conceptos de varios vocabularios compartidos.
- Una red semántica que especifica un conjunto de tipos semánticos que se aplican a los términos especificados en el metatesauro. También se especifican una serie de vínculos que se establecen entre los tipos semánticos.
- Un léxico especializado, que contiene términos de dominio general y del dominio de la biomedicina.

Otras ontologías que incorpora UMLS son la ya mencionada MeSH, LOINC (Códigos Identificadores de Observación Lógica), MEDLINE (Análisis de la literatura médica y Sistemas de Recuperación de Datos), UK Clinical Terms, entre otras. Una de las que goza de mayor aceptación es SNOMED-CT (Nomenclatura Sistematizada de Medicina – Términos Clínicos), que provee identificadores para conceptos y para los términos que representan a esos conceptos, sinónimos en múltiples idiomas, mecanismos de localización de preferencias, extensión con términos locales, posibilidad de definición de subconjuntos, etc. Los componentes básicos son los conceptos, que llevan asociados una serie de relaciones y de descripciones, como puede apreciarse en la Figura 2.2

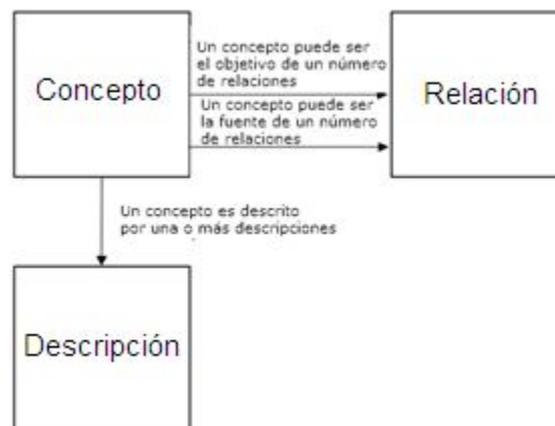


Figura 2.2. Núcleo de la estructura del SNOMED CT

## 2.4 Sistemas de soporte a la decisión clínica

Los sistemas basados en el conocimiento son el tipo más común de CDSS (Sistemas de Soporte a la Decisión Clínica). También denominados sistemas expertos, contienen conocimiento clínico sobre una tarea específica, y son capaces de obtener conclusiones a partir de datos relativos a un determinado paciente. Aunque existen muchas variaciones, el conocimiento dentro de un sistema experto está típicamente representado en forma de un conjunto de reglas.

Los CDSS toman un papel importante en el proceso de investigación científica y en particular, de la creación de nuevo conocimiento clínico. Por ejemplo, un computador puede ser utilizado para analizar una cantidad ingente de datos, buscando patrones complejos que sugieran asociaciones inesperadas. Igualmente, a partir de una lista de modelos computacionales, un sistema puede ser utilizado para observar un conjunto de observaciones experimentales que intentan coincidir con una serie de teorías dadas.

Este tipo de sistemas ayudan al médico a tomar las decisiones oportunas en el proceso de diagnóstico, a partir de una interacción activa entre diferentes modalidades médicas de conocimiento. Dentro de este contexto, pueden ser identificados diferentes modalidades [45], entre las que destacan: el conocimiento tácito de profesionales en términos de destrezas, juicios e intuiciones, experiencias clínicas (registradas y observadas), discusiones colaborativas o consultas entre practicantes, literatura médica publicada, contenido educacional médico para practicantes y pacientes, soporte para el conocimiento basado en observaciones clínicas, tests de diagnósticos y tratamientos terapéuticos, entre otros.

Más concretamente, en [34] introducen un concepto relacionado con este tema, denominado *knowledge morphing*, y que trata de integrar contextualmente, conceptualmente y funcionalmente, objetos de conocimiento que pueden existir en diferentes representaciones y formalismos. El objetivo es el de establecer una visión de múltiples facetas del conocimiento relativas a un problema específico dentro del contexto sanitario. Este concepto es altamente atrayente dentro de los CDSS debido a la existencia de múltiples modalidades de conocimiento médicas y la necesidad imperiosa de integrarlas para adquirir un entendimiento coherente a partir de todas las soluciones pertinentes, puntos de vistas, resultados disponibles, y en general cualquier fuente de información útil (Figura 2.3). Esta integración de múltiples fuentes de información de diversa naturaleza permite una evaluación del paciente más eficiente, e introduce el concepto de diagnóstico multimodal [35]. Otro trabajo interesante dentro de este contexto se presenta en [46], donde se expone un sistema de teleasistencia multimodal diseñado para extraer diferentes tipos de información, realizando un análisis por medio de un agente multimodal construido para analizar el estado de salud de la persona por medio de una fusión de toda la información que se extrae en un instante preciso.

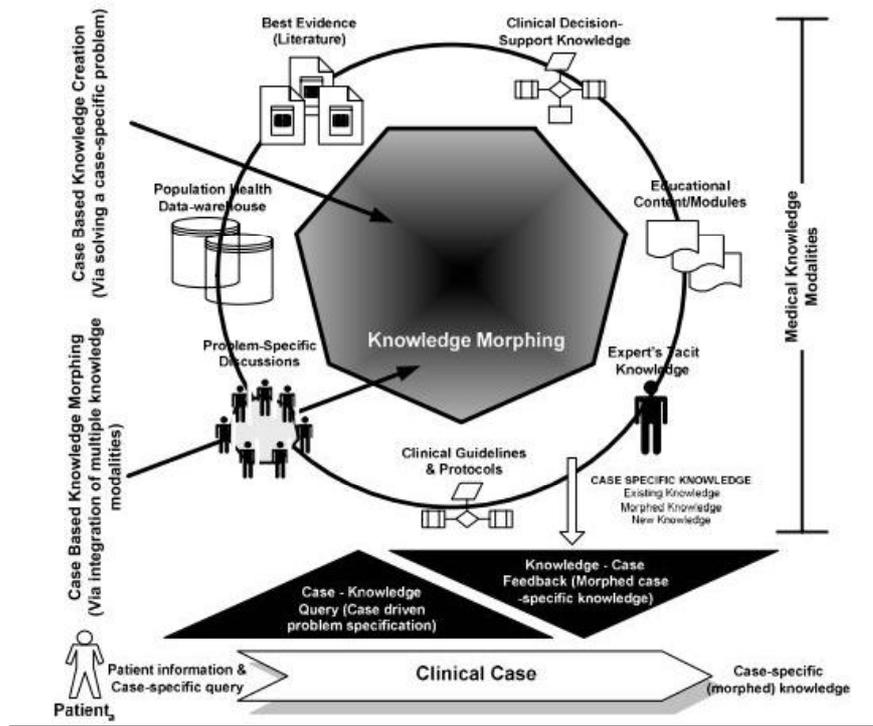


Figura 2.3. Un entorno de trabajo para la gestión del conocimiento

Los CDSS pueden clasificarse en función de la tarea de la que se encargan:

- Alertas y recordatorios: en situaciones de tiempo real, un sistema experto monitoriza las constantes vitales de un paciente mediante un electrocardiograma por ejemplo, y puede enviar alertas a una determina pantalla o vía correo electrónico para notificar al personal pertinente acerca de los resultados, o sobre tareas importantes que deben llevarse a cabo antes de que un evento ocurra.
- Asistencia al diagnóstico: los CDSS puede ayudar al personal sanitario pertinente a formular un diagnóstico más eficiente gracias al conocimiento que proporcionan.
- Prescripción de medicamentos: una de las tareas más comunes dentro del ámbito sanitario es la prescripción de medicamentos, por lo que un sistema experto puede automatizar este proceso y ofrecer otros servicios, tales como la generación de un mensaje electrónico que informe a la farmacia más cercana.
- Reconocimiento de imágenes e interpretación: los CDSS pueden servir de apoyo a la hora de interpretar imágenes médicas de diferente índole, tales como imágenes de rayos X, angiogramas o escáneres MRI.

## 2.5 Modelado multi-escala matemático

La generación de conocimiento por parte de los modelos computacionales multi-escala, y su uso en la práctica clínica diaria, es potencialmente clave para conseguir una mayor personalización de la asistencia sanitaria, así como para obtener conocimiento aplicable en el cuidado integral de la salud de una persona, ayudándole en las tareas de prevención, diagnóstico, terapia, y seguimiento de la evolución.

La importancia del modelado multi-escala resultará vital en un futuro, puesto que actualmente la investigación científica se centra en analizar el cuerpo humano a partir de la abstracción en componentes más sencillos. Sin embargo, para concebir una visión más amplia, es necesario un nuevo enfoque que trate de interconectar las piezas e intente inferir una relación o significado entre las mismas.

Este desarrollo debe llevarse a cabo por medio de la unificación de componentes en función de sus escalas temporales y dimensionales, y por medio de la integración de disciplinas tales como la bioingeniería, la medicina o la biología. La cantidad ingente de información ya existente, y la nueva que surge con los avances científicos, hacen clave el concepto de sistema gestor de conocimiento.

Actualmente existen dos paradigmas de computación multi-escala [36]. El primero se basa en un acoplamiento débil, que viene ilustrado en la Figura 2.4, donde un modelo de una determinada escala puede proveer como entrada a otro modelo de otra escala distinta. Los modelos son ejecutados una sola vez, y el flujo de información discurre de un elemento a otro.

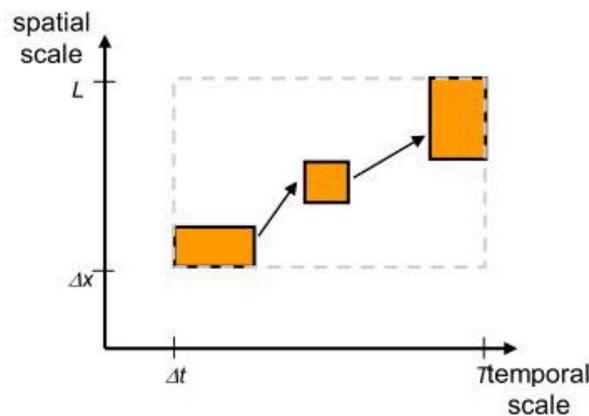


Figura 2.4. Acoplamiento débil en computación multi-escala

En cambio, el paradigma de computación multi-escala de acoplamiento fuerte sigue un bucle iterativo donde los modelos pueden ejecutarse varias veces, se encuentran estrechamente unidos, y son necesarias librerías de acoplamiento para llevar a cabo la implementación. Este paradigma viene ilustrado en la Figura 2.5.

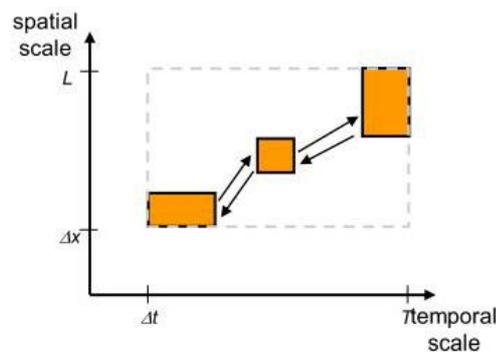


Figura 2.5. Acoplamiento fuerte en computación multi-escala

Dentro de la informática biomédica, el área de la computación distribuida resulta un campo sumamente atractivo para comunicar diferentes modelos localizados en lugares dispares y con diferentes tecnologías. En este contexto, el modelo de comunicación basado en el paradigma publicador/suscriptor resulta uno de los más idóneos para la comunicación de información entre modelos. En [3] se realiza un primer acercamiento siguiendo este enfoque mediante el uso de la especificación DDS como sistema de distribución de datos entre modelos multi-escala, proveyendo además de la característica deseable de tiempo real.

Ligado al modelado multi-escala, se encuentra el concepto de personalización sanitaria. En [37] desarrollan el uso de un sistema que mide la pulsación arterial en tiempo real utilizando las propiedades ópticas de interacción de la luz con la materia, y hace uso de una representación fisiológica multi-escala para investigar simultáneamente procesos en diferentes niveles jerárquicos, aportando una nueva perspectiva para la personalización clínica del paciente.

Otros estudios han intentado abordar esta problemática mediante el diseño de procedimientos de diagnóstico que simulan el comportamiento del estado interno de los pacientes, utilizando modelos matemáticos multi-escala. Los ejemplos incluyen proyectos tales como ERATO Systems Biology Workbench [38], cuyo objetivo es crear un marco computacional que permita una comunicación para compartir datos y algoritmos de distintos componentes escritos en diversos lenguajes de programación y corriendo en plataformas dispares.

Otro trabajo destacable se expone en [39], donde se presenta una metodología de generación de conocimiento dentro del contexto del modelado multi-escala, y que se encuentra basado en tres conceptos claves: la inteligencia de la capa de sensores, la generación de información del estado interno del paciente mediante modelos matemáticos que se adaptan a la evolución del paciente por medio de variables monitorizadas, y la descripción mediante componentes de simulación modulares y reutilizables pertenecientes a diferentes escalas.