

Capítulo 4: Análisis funcional del sistema

4.1 Características y especificaciones de diseño

El modelo de datos empleado en este nuevo diseño para la compartición de recursos está basado en la distribución de elementos constituyentes de un registro mecanizado que cubre los datos sociales, preventivos y médicos de un determinado paciente, y que recibe la denominación de Historia Clínica Electrónica (HCE). En general, estos datos son almacenados en repositorios compatibles que siguen, por un lado, un determinado modelo de referencia que indica cómo debe presentarse o almacenarse la información en soporte informático, y por otro, un modelo de arquetipos, que define un concepto clínico determinado mediante restricciones en el modelo de referencia. De esta forma se hace posible la construcción de compartimentos de información denominados extractos, para el almacenamiento de la historia clínica de un paciente. Normalmente esta información se encuentra localizada entre diversos sistemas autónomos y heterogéneos cuya interoperatividad se hace difícil de lograr [49].

Los intentos de desarrollo de estándares para la representación y comunicación de la información de la HCE han sido abundantes en los últimos años. Sin embargo, la complejidad de la información sanitaria, su heterogeneidad en cuanto a tipos de datos representados, y los continuos avances del conocimiento médico, repercuten en los modelos propuestos, quedándose obsoletos o afectando a su mantenimiento y gestión con el paso del tiempo [50].

Estos modelos definen descripciones de ideas abstractas, procedentes del mundo real o no, y que siguen una estructura formalizada para que los humanos puedan entenderla. En general, en el dominio de la medicina clínica, y particularizando en el campo de la HCE, se hace patente la dificultad para la obtención de modelos válidos, es decir, estándares eficientes para el desarrollo de una arquitectura que maneje extractos HCE, mensajes asociados y terminología definidas en el estándar pertinente.

Un nuevo enfoque que plantea una nueva visión a este problema, y que surge como propuesta para el diseño de sistemas de información sanitarios, está basado en un modelo dual [51] de desarrollo que permite realizar una distinción entre información y conocimiento, controlando los conceptos del dominio manejados por los futuros sistemas de información. Por un lado, la información reúne el conjunto de reglas para construir sentencias, o lo que es lo mismo, la definición de una gramática descrita por un modelo de referencia, y por otra parte, el conocimiento recogería modelos formales de un concepto clínico en uso, basándose en un modelo de referencia concreto y limitándolo por medio de restricciones. Este conocimiento vendría representado por un modelo de arquetipos.

Dentro de la arquitectura que se propondrá a continuación, es necesario definir varios tipos de roles o actores que se encargarán de interactuar con cada uno de los módulos del sistema, ya sea aprovechando las funcionalidades presentes, añadiendo nuevas, u ofreciendo soporte tecnológico:

- Usuario: se define como toda entidad física o lógica ajena al sistema con la capacidad de consumir y/o producir información o conocimiento, y que explota los servicios proporcionados por el sistema. Ejemplos de usuarios dentro del sistema pueden ser dispositivos portables con sensores y sin intervención manual, plataformas de adquisición o monitorización de datos, modelos matemáticos multi-escala, profesionales sanitarios o pacientes.
- Experto en el dominio: conocedor de la norma sanitaria implementada, y encargado de

realizar la definición de arquetipo, especificando una serie de características y estructuras que deben seguir la representación formal de un determinado concepto clínico y que está basado en un modelo de referencia específico.

- Experto tecnológico: conocedor de la arquitectura del sistema al completo, de la estructura del modelo de información y del modelo de arquetipos. Se encarga de la implementación de estos modelos, de sus interfaces, y de los mapeos entre los arquetipos y las fuentes de datos, además de ofrecer soporte tecnológico y de mantenimiento.

El diseño del sistema gestor de información basado en el conocimiento se realiza mediante una descomposición modular de cada una de sus partes, posibilitando la ampliación y modificación de nuevas funcionalidades y nuevos módulos.

Todos los componentes del sistema pueden encuadrarse dentro de cuatro partes bien diferenciadas:

- Bloque de Gestión del Conocimiento (CGB): gestiona la creación, transmisión, almacenamiento y visualización de información y del conocimiento, agrupando cada una de estos procesos en módulos, y abstrayendo cada una de las partes para una mejor descripción de sus componentes constituyentes.
- Bloque de Administración del Sistema (SAB): controla la gestión, el mantenimiento y la recuperación de errores dentro del sistema. Se encuentra íntimamente vinculado al anterior, y posee autoridad absoluta sobre todos los módulos del sistema, pudiendo habilitar o deshabilitar cualquiera de ellos en un determinado momento. También se encarga de la gestión del arranque del sistema de cada uno de los componentes dentro de la arquitectura, de la administración del clúster de base de datos dentro del módulo de almacenamiento de información y conocimiento, del mapeo de los tipos de datos utilizados en el middleware y los tipos de datos empleados en la base datos objeto-relacional
- Middleware Orientado a Mensajería (MOM): es el centro neurálgico de la arquitectura y encargado de la transmisión y distribución de los datos entre todos los módulos del sistema. Sigue el paradigma de comunicación publicador/suscriptor, que presenta un conjunto de ventajas respecto el modelo cliente-servidor, permitiendo un mejor desacoplamiento entre nodos, una tolerancia a fallos más depurada, una menor latencia, y un sistema de conexión entre nodos intrínsecamente más sencillo. Como se verá posteriormente, la elección del Middleware se considera crucial, ya que según las funcionalidades que ofrezca, el sistema gestor podrá ofrecer unos servicios u otros.
- Interfaz de usuario: es el medio de interacción gráfico que poseen los diferentes usuarios que pueden actuar sobre el sistema. Debido a la gran diversidad y privilegios de los mismos, es necesario definir una política de credenciales suficientemente robusta como para proveer un acceso adecuado a la información y al conocimiento. También es imprescindible la implementación de un mecanismo de auto-aprendizaje que permita el almacenamiento dentro de un perfil de usuario, de los servicios utilizados habitualmente por el mismo, personificando y adaptándolos en función de las necesidades y de la práctica clínica.

En la Figura 4.1 se ilustra cada una de las partes comentadas anteriormente dentro del sistema propuesto:

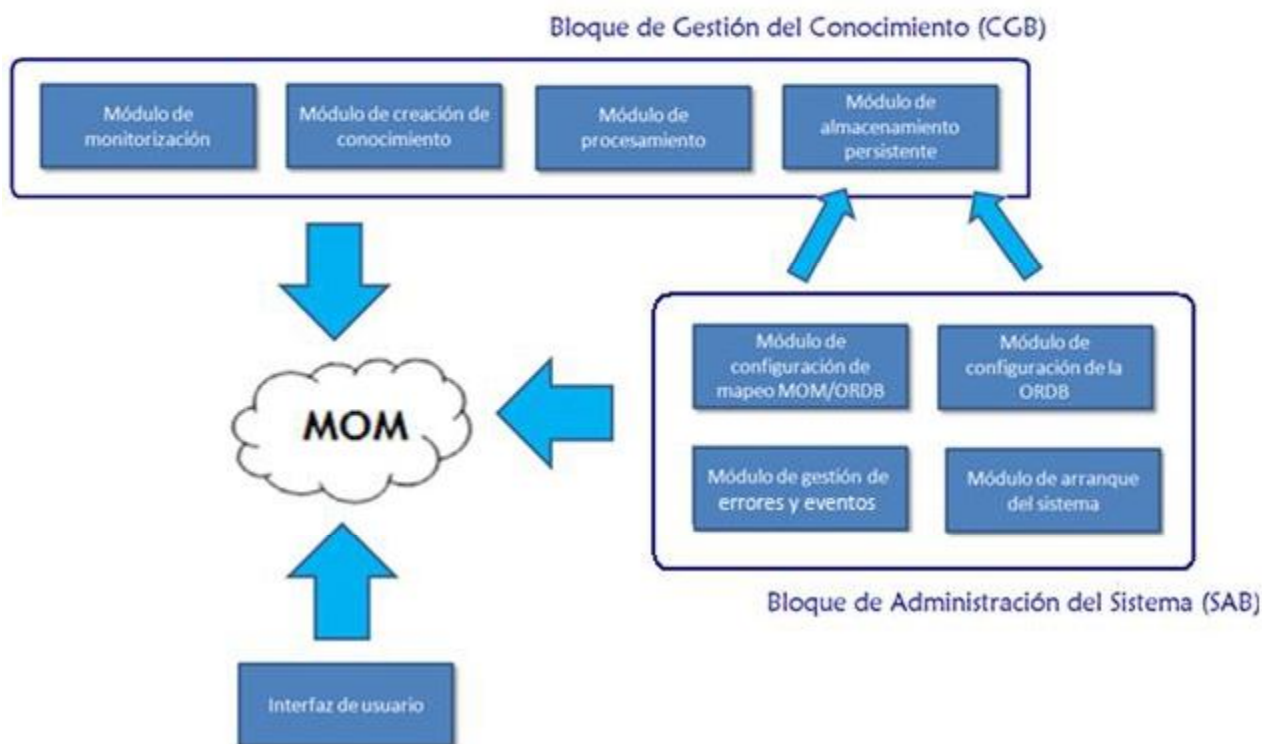


Figura 4.1 Nuevo Diseño del Sistema gestor de datos en tiempo real basado en el conocimiento

4.2 Descripción modular del Bloque de Gestión del Conocimiento (BGC)

4.2.1 Introducción

La abstracción de funcionalidades dentro del BGC se efectúa en base a la separación de métodos o técnicas a partir de los términos de información y conocimiento. El concepto de datos no entra dentro de esta discriminación, ya que al tratarse de unidades mínimas semánticas que se corresponden con elementos primarios de información, son irrelevantes como apoyo a la toma de decisiones.

Las técnicas consideradas para la gestión de la información y del conocimiento son las siguientes:

- **Adquisición de información:** se encarga de la toma de datos por parte de elementos externos al sistema, ya sean medidas del estado de salud del paciente tomadas en tiempo real (pruebas diagnósticas, monitorizaciones, exploraciones,...) o no (analíticas, informes sanitarios,...), y en general, de cualquier entidad lógica o física con la capacidad de producir información. Esta obtención de información se origina fuera del sistema, por lo que dentro del mismo no resulta conveniente hablar de creación o generación, sino de adquisición, ya que los datos se están tomando de entidades que no pertenecen al sistema.
- **Creación de conocimiento:** cuando se habla de conocimiento, es necesario contemplarlo desde otro punto de vista. En este caso, el conocimiento no se adquiere, sino que se crea a partir de la información o del conocimiento ya existente. En el caso concreto de nuestro

sistema, esta creación va a venir dada por dos pasos:

- La especificación de conceptos clínicos que vienen representados por instancias de un modelo de conocimiento o arquetipos. Estas instancias son interpretaciones de entidades de un modelo de referencia en concreto, y se establecen bajo una serie de restricciones que delimitan su comportamiento. Son elaboradas por expertos en el dominio conocedores de la norma sanitaria con la que se trabaja, y da como resultado una plantilla compuesta por extractos de un modelo de arquetipos definido. Ejemplos de conceptos clínicos pueden ser la temperatura, la presión arterial, la concentración de glucosa en sangre, entre otros.
- Trata de la generación propiamente dicha de conocimiento, utilizando las plantillas de datos constituidas en el paso anterior. En esta generación de conocimiento van a estar involucrados todo tipo de actores capaces de esta generación.
- **Procesamiento de información y conocimiento:** en donde se utilizan mecanismos automáticos que generan nueva información y conocimiento a partir del ya existente. Uno de los principales elementos involucrados en esta generación son los modelos matemáticos computacionales multi-escala, que puedan estar ejecutándose en tiempo real y alimentándose de diferentes fuentes de datos dispares e incluso de resultados de otros modelos corriendo en ese momento.

Para que este procesamiento y generación de nuevo conocimiento se produzca adecuadamente, es necesario establecer las entradas de datos del modelo, es decir, las características de los datos de entrada al mismo, así como los resultados producidos después del proceso. Computacionalmente hablando, el objetivo consistiría en definir unas interfaces de entrada y unas interfaces de salida que puedan ser utilizados por otros elementos de comunicación.

Por otro lado, se considera importante también mantener la autonomía de cada uno de los sistemas de modelado y utilizar los procedimientos apropiados para alcanzar una interacción flexible entre ellos, consiguiendo por tanto definir un escenario de computación de modelos distribuido.
- **Almacenamiento de información y conocimiento:** controla que tanto la información y el conocimiento sean guardados en lugar pertinente y en el momento apropiado (en el caso en el que trabajemos en tiempo real).

Para la información o el conocimiento que deba almacenarse de forma no volátil, debe proporcionarse un conjunto de bases de datos objeto-relacionales que deben seguir una metodología propia de diseño de bases de datos distribuidas. Hay que tener en cuenta que, para el resto de módulos que componen el sistema, la topología de bases de datos distribuidas diseñada es irrelevante en este sentido, por lo que debe garantizarse transparencia en el acceso a los datos.

Para la información y el conocimiento volátil (la que sólo permanece en el sistema durante un periodo de tiempo definido), o aquella que utilizan determinados procesos de forma frecuente, el MOM debe proveer un segmento de almacenamiento en memoria principal para distribuir los recursos disponibles de la forma más eficiente posible. Igualmente, todos los mecanismos de sincronización, exclusión mutua y demás aspectos relacionados con el acceso concurrente a los datos debe suministrarlo el MOM de forma transparente, de la misma forma a como lo efectúa la base de datos objeto-relacional con los datos almacenados de forma no volátil. Si en cambio el MOM no proporciona esta funcionalidad, se opta por la construcción de un mecanismo de conexión de bases de datos in-memory.
- **Visualización de información y conocimiento:** que trata de exponer a los distintos usuarios de la manera más entendible posible, la información o el conocimiento almacenado dentro del sistema. En este apartado entraría en juego el desarrollo de una interfaz de usuario intuitivo, acorde con las necesidades exigibles. Debido a la gran diversidad de usuarios que pueden hacer uso de los servicios proporcionados por el sistema, resulta una tarea ardua la construcción de una interfaz genérica para todos ellos, por lo que se hace necesario una

distinción en función de los roles de los usuarios, un análisis pormenorizado de los requerimientos y privilegios para cada uno de ellos, una especificación de control de acceso pertinente, y una política de seguridad suficientemente robusta. Esta disciplina no se ha contemplado dentro del Bloque de Gestión del Conocimiento, puesto que se ha considerado que también está relacionado con el Bloque de Administración del Sistema y por tanto se ha diseñado como un módulo ajeno a ambos bloques.

- **Transmisión de la información:** en donde la información y el conocimiento no son vistos como tal, sino como un flujo constante de datos, sin considerar el formato o la semántica de los mismos. Esta técnica trata sobre el transporte de datos de un elemento de comunicación a otro o a otros, y debe apoyarse en una tecnología que permita ubicar dichos datos en el lugar correcto y en el momento adecuado.

4.2.2 Módulo de monitorización

Controla la obtención de información por parte de fuentes heterogéneas productoras de datos externas al sistema, con el objetivo de adquirir medidas de las constantes vitales (presión sanguínea, frecuencia cardíaca) y otras señales biomédicas (ECG o ECO) de un determinado paciente, con el fin de enviarlas posteriormente a un almacén de datos donde puedan estar disponibles para la consulta del profesional pertinente.

Dentro de los dispositivos de adquisición y medición de parámetros vitales, los avances en el campo de la Ingeniería Biomédica han dado lugar al desarrollo de equipos cada vez más precisos, manejables, accesibles para el paciente, e integrados completamente en la actividad cotidiana gracias al surgimiento de los dispositivos portables.

De esta forma, la monitorización de las constantes vitales del paciente puede llevarse a cabo de forma domiciliaria (el paciente transmite la información desde su propio domicilio), ambulatoria (donde se hace uso de dispositivos de monitorización dentro de entornos controlados, como hospitales), o ubicua (donde la provisión de la información es completamente independiente de la localización del paciente). Sin embargo, a día de hoy existe cierta heterogeneidad entre estos dispositivos que no las hacen integrables con otras aplicaciones similares en el contexto sanitario global. En la actualidad, un aspecto clave a tener en cuenta en este tipo de sistemas, es el diseño basado en normas que definan completamente las reglas con las que dispositivos médicos se conectan al sistema, encapsulan los datos y los transmiten.

El estándar para la comunicación de dispositivos médicos dentro de la Unión Europea es la norma ISO/IEEE 11073 (X73), que proporciona la versatilidad suficiente para convertir la información en un formato interoperable para que pueda ser utilizada por otro tipo de dispositivos o por sistemas. Estos datos pueden formalizarse en componentes que constituyan extractos de HCE, construyendo sistemas más genéricos que provea soluciones globales. Actualmente existen trabajos que tratan de afrontar este desafío, integrando datos transmitidos desde dispositivos médicos según la norma X73 sobre un servidor de historias clínicas electrónicas, a través de los interfaces de comunicación que incorporen (Wireless, ZigBee, Bluetooth,...).

En nuestro caso concreto, el objetivo del módulo de adquisición de información va a ser el de mapear los datos provenientes de cualquier dispositivo de monitorización, a estructuras constituyentes de un extracto de HCE.

La información, independientemente de cómo sea presentada, es decir, en qué formato sea entregada al módulo de adquisición de información a través de una de sus interfaces, debe ir acompañada de un fichero de configuración que permita el mapeo de estos datos a un metalenguaje

ya estandarizado que facilite el tratamiento de esta información dentro del sistema gestor. Con esto se pretende facilitar la integración de los datos, dándole un contenido semántico a los mismos. Esta fase se antoja la más complicada de superar, ya que el archivo de configuración debe estar correctamente construido para que el mapeo de datos se produzca de forma satisfactoria.

Una vez alcanzado este punto, el contenido de la información ya estandarizado debe convertirse a un extracto de HCE para favorecer un posterior tratamiento de la información. Para llevar a cabo este proceso, se presenta un componente de transformación conforme al metalenguaje estandarizado comentado con anterioridad, y que trata de introducir los datos dentro de un extracto específico del modelo de arquetipos que sigue una norma sanitaria concreta

Por último, este extracto debe ser publicado dentro del contenedor de información propio del middleware empleado en el sistema. Este trasvase de datos puede realizarse de dos formas:

- Traspase en “bruto”: donde la información es publicada en el middleware sin que se lleve a cabo ningún tipo de validación, y de forma indivisible, por lo que el suscriptor recibirá todo el extracto de forma íntegra, siendo él el encargado de distinguir entre la información relevante y la prescindible.
- Traspase normal: en este caso, cada uno de los campos correspondientes del extracto de HCE debe mapearse a los tipos de datos empleados por el sistema middleware. Se hace por tanto necesario la existencia de algún tipo de mecanismo explícito que permita esta conversión. Este método se emplearía cuando el publicador conoce de antemano qué parte de la información contenida dentro del extracto completo es la que le interesa al suscriptor o suscriptores, realizando él la discriminación de los datos, repercutiendo en una disminución del ancho de banda utilizado, y relegando de esta labor al suscriptor. De esta forma se permite un filtrado de información más personalizado.

En la Figura 4.2 se muestra una ilustración para una mejor comprensión de lo comentado con anterioridad:

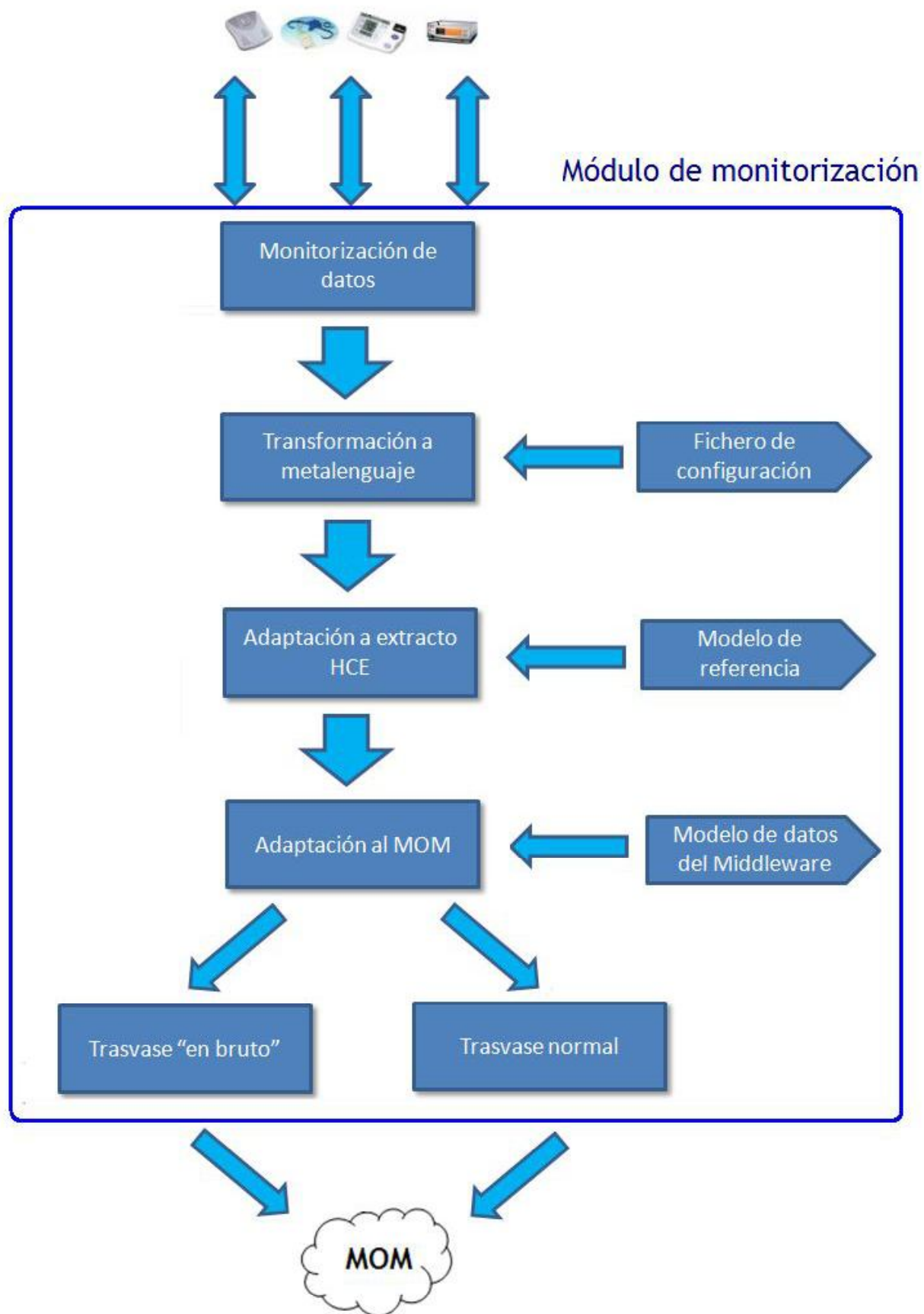


Figura 4.2: Módulo de monitorización

4.2.3 Módulo de procesamiento

El éxito de los modelos matemáticos que se construyen en la actualidad dependen en gran medida del conocimiento que éstos sean capaces de generar y de su capacidad de predicción. Normalmente los resultados van a depender en gran medida de la adquisición de información desde otros elementos de comunicación involucrados y que trabajan bajo unas fuertes restricciones de tiempo real.

La construcción de estos modelos computacionales multi-escala que tratan de describir la dinámica de los sistemas fisiológicos se presenta como una tarea compleja. Actualmente se utiliza una metodología jerarquizada que trata de representar los comportamientos del sistema a diferentes escalas, cuya integración total producirá el comportamiento global de todo el modelo. Esta aproximación jerárquica también reduce el riesgo asociado a la simplificación fisiológica necesaria para la construcción del modelo. Los distintos niveles implicados describen fenómenos a escalas temporales y espaciales muy diversas, siendo necesaria una integración adecuada para extraer conocimiento relevante.

El análisis de funciones biológicas de integración, y en general, la extracción de conocimiento en términos de estructuras subyacentes y mecanismos moleculares, es a día de hoy un aspecto clave en el modelado matemático de sistemas biológicos. Esta extracción de conocimiento de fuentes de datos heterogéneas debe producirse a partir de interfaces de entrada y de salida correctamente definidos. Estas interfaces deben proporcionar interoperatividad entre plataformas dispares a todos los niveles para que la integración entre modelos sea total.

Particularizando dentro del contexto en el que nos encontramos, estas interfaces van a venir dadas por un modelo de datos dinámico que va a suministrar el MOM. El nexo de unión entre modelos, que sirve como elemento intermediario para la producción o consumo de datos, es el canal, mensaje, evento o topic (dependiendo de la tecnología middleware utilizada), basándose en el paradigma de comunicación publicación/suscripción.

Para lograr la unificación entre diferentes modelos debe proporcionarse a los mismos los datos de entrada necesarios para su ejecución. Para ello deben definirse un conjunto de canales de información de suscripción donde pueda obtenerse la información y/o conocimiento pertinente. Asimismo, cada modelo publicará sus resultados en otro canal o canales destinados a dicho fin. Entre los datos que deben manejarse, deben considerarse los componentes que permitan suministrar a los modelos los datos precisos sobre el paciente, en el momento y lugar adecuados (estos podrían a su vez ser resultados de otros modelos) y, finalmente, los elementos que reciben y explotan el conocimiento generado por los mismos.

Los modelos podrán adaptarse a la publicación de sus resultados en función de la demanda de los suscriptores, e incluso no generar ningún tipo de información si no existe en esos momentos suscriptores asociados o modelos asociados a ese canal de salida. Esto facilitaría la transmisión de datos entre modelos multi-escala, que es otro de los fines que se persiguen.

Los consumidores o suscriptores de información y conocimiento reciben automáticamente las actualizaciones de los datos que se publican, sin preocuparse de su origen. Se consigue de esta forma que los datos puedan proceder de distintas bases de datos, históricos, otros modelos computacionales, historia clínica... De la misma forma, los publicadores tampoco tienen que preocuparse del destino de la información o conocimiento que producen. Cuando los suscriptores se desconectan, dejan de recibir esos mensajes, y los pierden. Esto conlleva un importante ahorro de comunicaciones, ya que sólo se transmite lo indispensable y durante el tiempo necesario.

Durante el proceso de construcción del modelo hasta su puesta en ejecución dentro del sistema gestor, se distinguen varias etapas que se describen a continuación, y se ilustran posteriormente en la figura 4.3:

1. **Implementación del modelo:** las dinámicas de muchos sistemas se pueden describir en términos de ecuaciones diferenciales que se obtienen utilizando leyes físicas que actúan sobre el sistema. El comportamiento de un sistema dinámico a una entrada puede obtenerse si se resuelven las ecuaciones diferenciales que modelan dicho sistema.
El objetivo de esta etapa es la de emplear algún tipo de formulismo matemático para expresar relaciones, propósitos, variables, entidades u operaciones, para estudiar los comportamientos dinámicos de sistemas fisiológicos complejos ante situaciones difíciles de observar en la realidad. Para ello, es necesaria la utilización de alguna herramienta matemática y lenguaje de modelado de sistemas dinámicos destinados a este propósito.
2. **Especificación de interfaces de entrada y de salida:** en donde se describen los puntos de comunicación de entrada y de salida del modelo construido. Estos vienen representados como canales o eventos dentro del MOM y pueden encontrarse establecidos de antemano o no. En caso de que no lo estén (o de que sí lo estén pero no se adapten a lo demandable o lo ofrecido por el modelo), se crearán los canales pertinentes dentro del middleware.
Estos canales son vínculos de unión entre modelos y generalmente contienen las variables asociadas a un determinado concepto clínico. Por tanto, estos canales pueden utilizarse como entradas, como salidas, o como ambas al mismo tiempo para un determinado modelo. La flexibilidad de adaptación a estos canales y las ventajas proporcionadas por el paradigma publicación/suscripción permite complejas asociaciones entre modelos, de tal forma que un determinado cambio en de una variable de entrada permite, por ejemplo, que varios modelos suscritos al canal involucrado en esta alteración puedan activarse, publicando sus resultados en otros canales de salida específicos, y posibilitando por tanto la ejecución de otros modelos que se encuentran interconectados a través de estos canales de salida. Este proceso podría continuar sucesivamente o retornar a un canal anterior facilitando la retroalimentación entre modelos.
3. **Configuración de los parámetros de comunicación:** los vínculos entre modelos y canales del MOM están relacionados por una serie de condiciones de comunicación que normalmente vienen dadas a través de unas políticas de calidad de servicio (QoS). Este servicio debe proveerlo el middleware, y en caso de que no lo haga, debe implementarse por separado.
En esta etapa entraría en juego el concepto de tiempo real reclamado por algunos modelos en particular. En este caso, deben proveerse mecanismos para entregar muestras de información a estos modelos dentro de unos intervalos de tiempo bien definidos. En estos modelos, su propia dinámica obliga a alimentarse de muestras de información bajo unas fuertes restricciones temporales. Una muestra no válida o antigua puede afectar de forma negativa a los resultados producidos por dicho modelo.
En esta parametrización de la comunicación, también deberían contemplarse aspectos relacionados con el control de flujo, la entrega ordenada de la información, las transmisiones confiables o no, la latencia máxima permitida, o la prioridad de la información, entre otros. A todo esto hay añadir la necesidad definir políticas que resuelvan conflictos cuando varios modelos publiquen información en un mismo canal.
4. **Ejecución del modelo:** una vez superados las fases anteriores, el modelo está preparado para ser utilizado por cualquier usuario que participe de los servicios del sistema. Su disponibilidad es publicada en un canal destinado a dicho fin, de tal que forma que los usuarios interesados reciben el anuncio de forma inmediata y puedan comenzar a ejecutar el modelo en cuestión.

Módulo de procesamiento

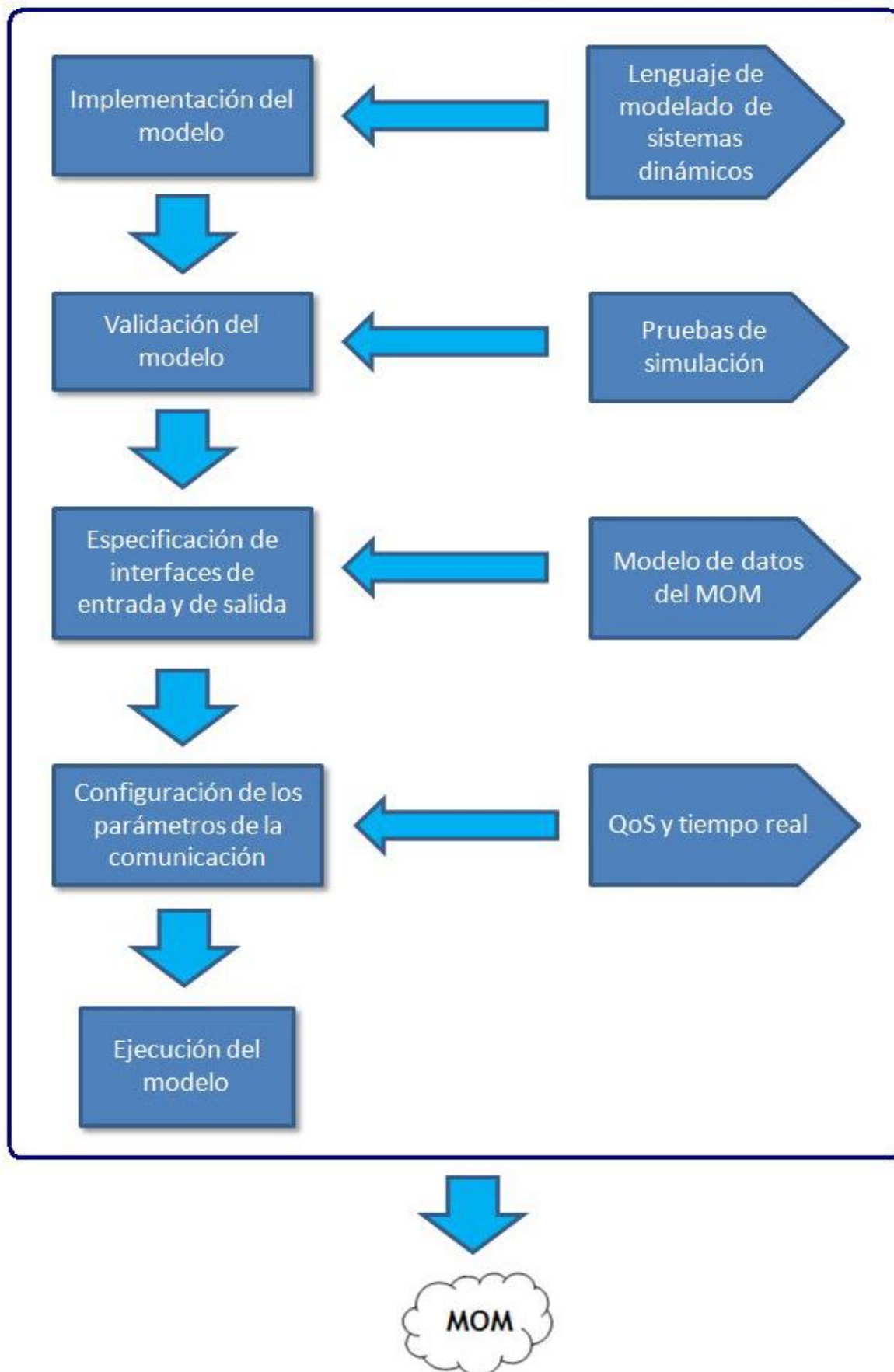


Figura 4.3: Módulo de procesamiento

4.2.4 Módulo de creación de conocimiento

La propiedad de interoperatividad dentro de un sistema de información sanitario distribuido puede verse afectada drásticamente si éste carece de una definición semántica adecuada de la información que maneja. Debido a la complejidad y tamaño del dominio clínico, la solución a este problema no se considera trivial. Existe una gran cantidad de conceptos clínicos que nunca estarán totalmente definidos debido a su constante evolución. A todo esto hay que sumarle que diferentes clínicas y hospitales tienen sus propios sistemas de información y de mantenimiento, por lo que la cantidad de recursos heterogéneos se multiplica.

El HCE tiene como objetivo proporcionar una medida estandarizada para afrontar esta adversidad e incorporar interoperatividad semántica. La finalidad de este módulo es la formalización de arquetipos acorde a un modelo basado en el conocimiento para la especificación de extractos HCE, de tal manera que puedan utilizarse después por el personal sanitario. Esta definición de extractos está basada en el paradigma dual del modelo de datos, que realiza una distinción entre información y conocimiento.

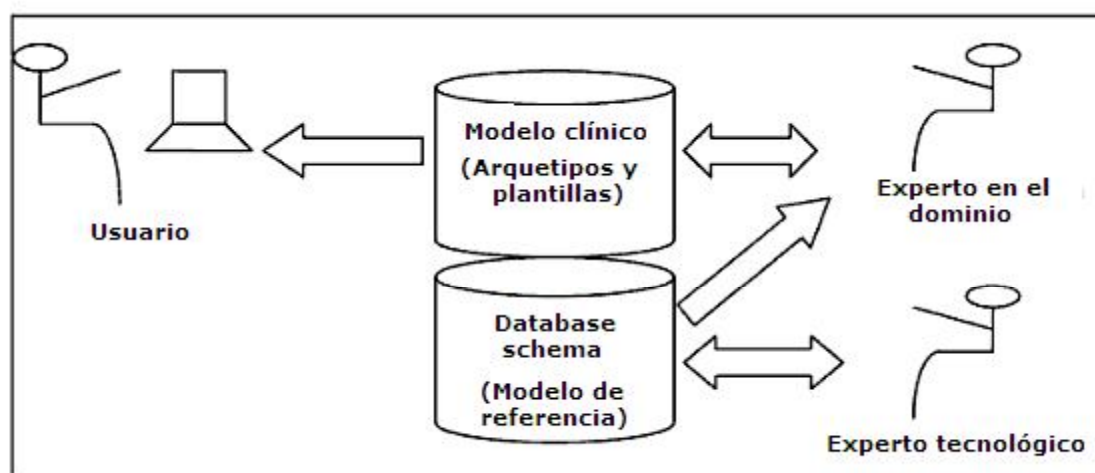


Figura 4.4: Modelo de referencia, arquetipos y actores involucrados

Este proceso se observa en la Figura 4.4, y van a estar involucrados todos los actores mencionados en apartados anteriores. Por un lado, el experto tecnológico que es el encargado de la implementación de las entidades correspondientes al modelo de referencia utilizado. A su vez, este modelo va a ser empleado por el experto en el dominio sanitario para la construcción de arquetipos y plantillas que posteriormente podrán ser accesibles para los usuarios de cualquier tipo.

En general, este desarrollo puede resumirse en la Figura 4.5.

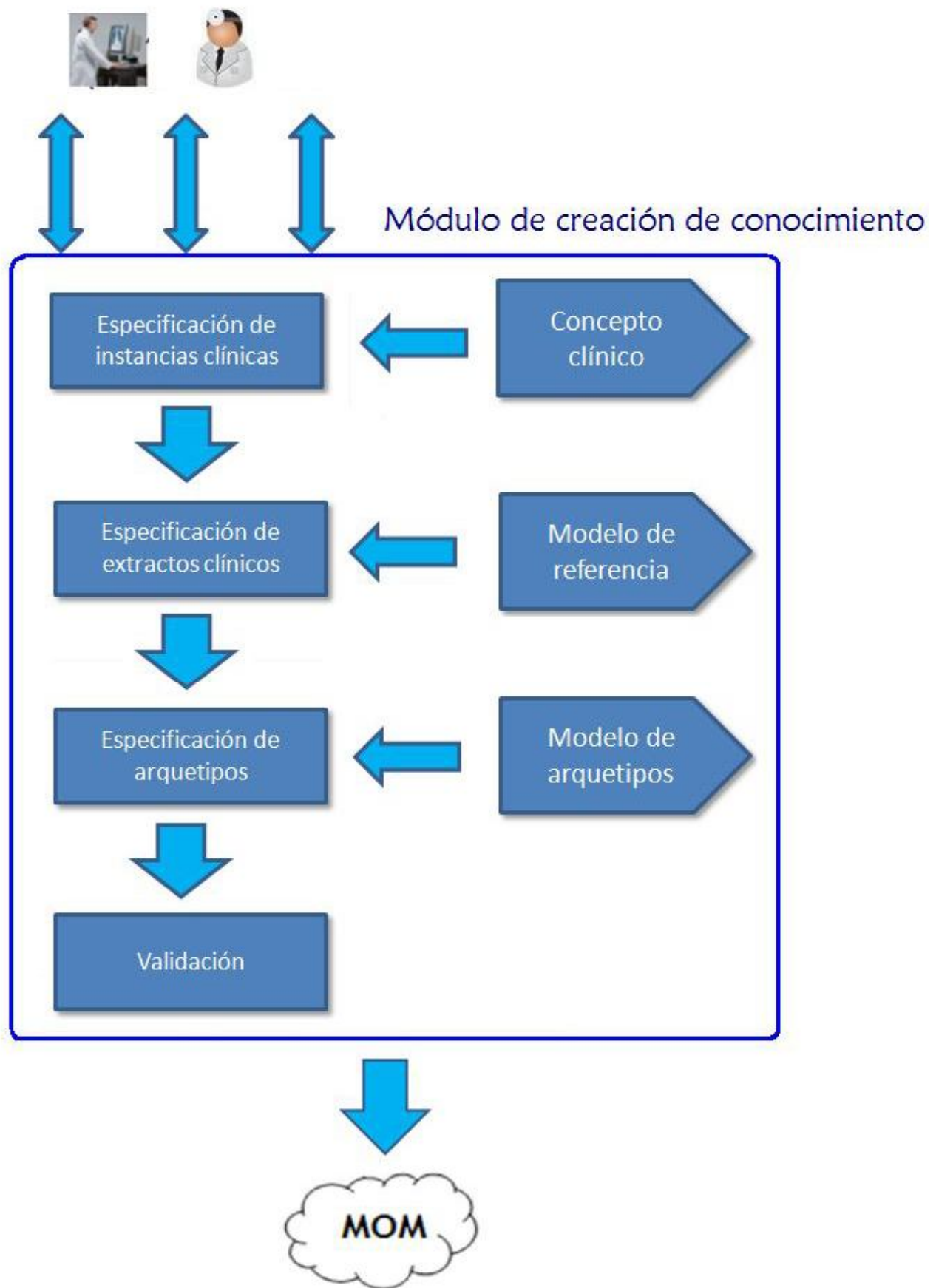


Figura 4.5: Módulo de creación de conocimiento

En primer lugar es necesario establecer qué concepto clínico es el que se va a modelar como arquetipo. Ejemplos pueden ser la presión sanguínea, la temperatura, la masa corporal, etc. A continuación es necesario fijar el contenido de este concepto clínico, que puede estar estructurado, semi-estructurado, no-estructurado o una combinación de todas, pudiendo además estar expresados en texto plano, texto de código, párrafo, medidas cuantificables con valores y unidades, fecha, tiempo, datos encapsulados, información multimedia, tipos de datos elementales, y en general cualquier tipo de dato soportado por el modelo de referencia empleado.

La especificación de extractos clínicos debe describirse por medio de un metalenguaje que proporcione contenido semántico a los tipos de datos elegidos en el paso anterior. Normalmente un modelo de referencia está construido a partir de entidades de datos yuxtapuestas que proporcionan varios niveles de abstracción para distinguir diferentes formalidades clínicas. Por ejemplo, los datos correspondientes al paciente y al personal sanitario, motivos de la visita, diagnóstico sanitario, tratamiento; u otras más elementales, como las medidas tomadas durante la evaluación del paciente, y que vienen representadas por tipos de datos básicos como enteros o flotantes. Este tipo de procedimientos tienen que almacenarse en entidades del HCE antes de la construcción del arquetipo y debe ser un proceso transparente para el personal clínico.

La especificación de arquetipos clínicos consiste en la acotación de aspectos relacionados con los atributos y los tipos de datos utilizados en los extractos de HCE establecidos anteriormente. Esta delimitación va estar relacionada en restricciones asociadas con el rango de los atributos, de su existencia, de su cardinalidad (si el atributo es multivaluado o no), de las ocurrencias, entre otros factores.

El resultado final debe expresarse en un lenguaje de definición de arquetipos que soporte restricciones y que posteriormente puedan ser analizados por una herramienta de validación que certifique si el arquetipo se ha construido de manera correcta.

Por último, el modelo conformado es trasvasado al MOM y almacenado de forma persistente en el módulo de almacenamiento para un posterior uso, y de forma volátil dentro de un repositorio o canal dentro del MOM, de tal forma que pueda estar accesible a cualquier usuario del sistema con la simple suscripción al mismo. La gran ventaja que proporciona el paradigma publicación/suscripción es que los elementos suscritos a este repositorio van a recibir automáticamente los arquetipos interesados, por lo que no necesitan realizar ningún tipo de petición previa.

4.2.5 Módulo de almacenamiento persistente

Este módulo tiene como principal componente un conjunto de bases de datos objeto-relacional, que van permitir almacenar los datos de manera persistente. Para asegurar la consistencia de los datos, las bases de datos debe ser *ACID Compliant* [64], es decir, deben cumplir un conjunto de características que garanticen que las transacciones llevadas a cabo se efectúen de forma fiable. Estas características definen aspectos relacionados con la atomicidad, consistencia, aislamiento y durabilidad de los datos. En la actualidad, la mayoría de las actuales bases de datos presentes en el mercado proveen estas propiedades.

Asimismo, el diseño de interconexión entre bases de datos debe seguir una metodología común de bases de datos distribuidas. Esta tarea queda fuera del objetivo de este trabajo. Sin embargo, unas especificaciones de diseño mínimas que deben cumplirse, estableciendo un modelo inicial de conexión de bases de datos basados en clústers-nodo, y un conjunto de servidores que controlan las peticiones entrantes de manera equitativa gracias un sistema de balanceo de carga.

Con esto se logra que la caída de una determinada base de datos o de un servidor no repercuta posteriormente en la extracción, inserción o actualización de datos. Estos elementos están representados en la Figura 4.6.

A continuación, se especifican las características que deben poseer cada uno de los elementos que componen este módulo:

- **ORDB (Base de Datos Objeto-Relacional):** en la elección de la base de datos, deben considerarse los siguientes aspectos:
 - Debe seguir el modelo objeto-relacional.
 - Debido a la variedad de los datos almacenar, es necesario que ofrezca soporte para tipos de datos BLOB, que permita el almacenamiento de archivos de gran tamaño.
 - Debe soportar el lenguaje SQL.
 - Debe proveer soporte para el acceso estandarizado a bases de datos
 - Es recomendable que incorpore la característica de triggers o disparadores, para que el mapeo de datos de la base de datos al middleware sea más eficiente.
 - Es recomendable, pero no imprescindible, que sea multiplataforma, compatible con alguna licencia de software libre, y suministre soporte para un amplio surtido de lenguajes de programación.
- **Nodo-clúster:** es una entidad que agrupa a varias bases de datos idénticas. Se debe proporcionar un mecanismo de replicación, en el cual, una de ellas tomaría el papel de base de datos maestra, y el resto de esclavas. Con esto se evita perder la característica deseable de disponibilidad en el caso en el que la base de datos primaria fallara, mejorando el rendimiento general del acceso a los datos gracias a una disminución de la latencia. Por tanto, la cantidad de copias dentro de un nodo-clúster es igual al número de nodos que contiene el grupo menos uno. Estas copias se denominan réplicas.
- **Clúster:** agrupa a un número determinado de nodo-clúster que se encuentran interconectados unos con otros a partir una metodología propia de diseño de bases de datos distribuidas.
- **Servidores:** se trata de la unión de varios servidores que trabajan como si de uno sólo se tratase. Son los encargados entregar los datos almacenados en las bases de datos al MOM.
- **Sistema de balanceo de carga:** cuya función va a ser la de gestionar las solicitudes entrantes de inserción, actualización o consulta en los clúster de bases de datos, y distribución del tráfico de forma equitativa entre todos los servidores.
- **MOM:** es el middleware orientado a mensajería que sirve de nexo de unión con el resto de módulos del sistema. Debe incorporar una plataforma de mapeo que permita traducir los registros contenidos en las tablas, a los tipos de datos que maneje el middleware orientado a mensajería en sus eventos, temas o mensajes (y viceversa). En caso de no incorporarlo, es necesario su implementación por separado en el módulo de configuración de mapeo MOM/ORDB.

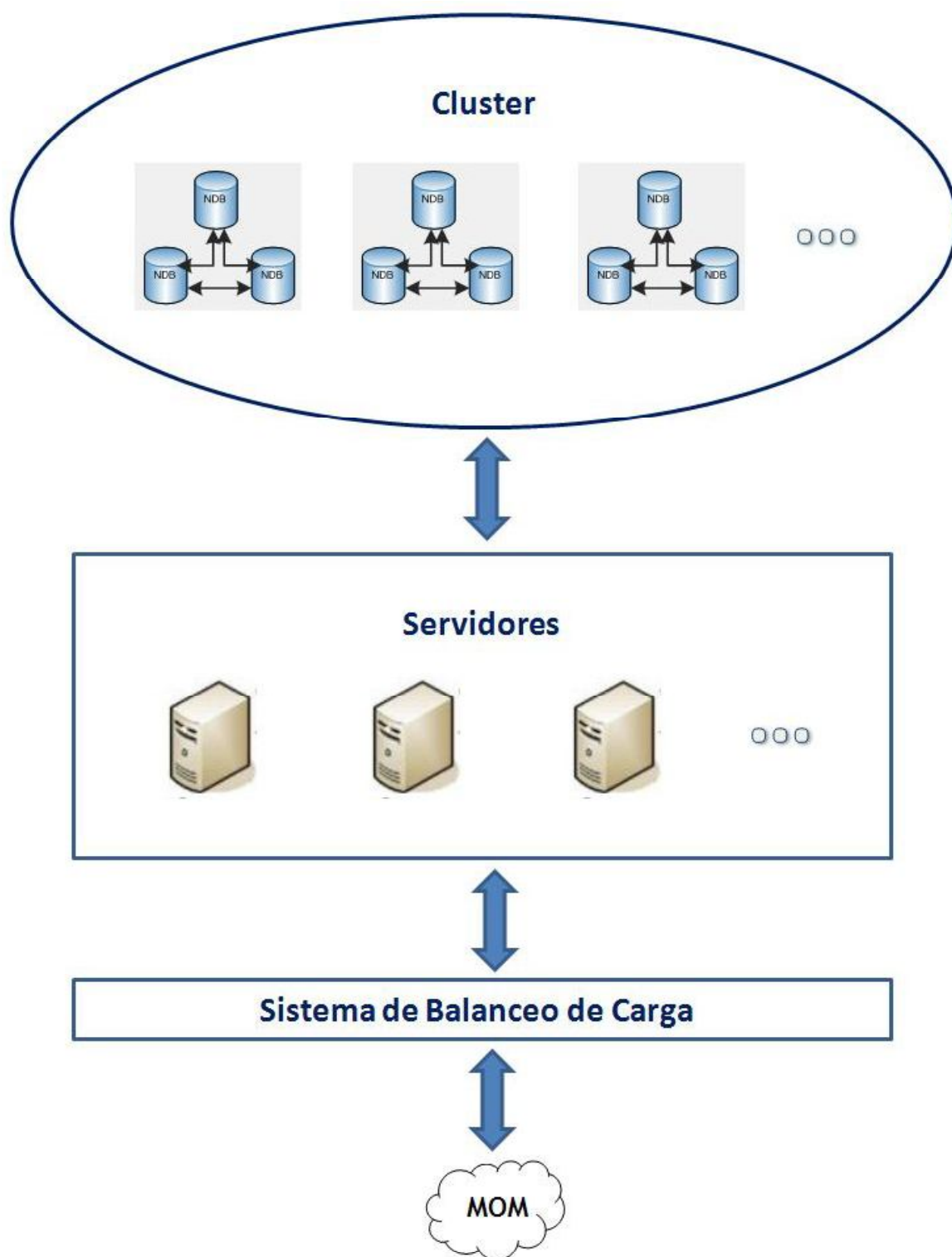


Figura 4.6: Módulo de almacenamiento persistente

4.3 Descripción modular del Bloque de Administración del Sistema (BAS)

El Bloque de Administración del Sistema (BAS) está fuertemente ligado al del Conocimiento (BGC) y depende de él. Actualmente, y debido al incipiente estado de los módulos del BAS, algunos módulos del BGC no se encuentran descritos a día de hoy o se encuentran poco desarrollados.

4.3.1 Módulo de arranque del sistema

El módulo de arranque del sistema es el que se encarga de ir habilitando cada una de las partes o módulos constituyentes del sistema. Se distinguen dos tipos de arranque: arranque inicial del sistema y arranque “en caliente” del sistema.

Arranque inicial del sistema

En el arranque inicial se establecen una serie de etapas que deben ejecutarse de forma secuencial, de manera que cada una de ellas no pueda iniciarse hasta que la anterior haya finalizado con éxito.

Si durante el transcurso de alguna de estas fases se produce algún error, debe notificarse a través del evento o mensaje definido para tal fin dentro del MOM. De esta forma, cualquier entidad o módulo puede percatarse de dicha contingencia mediante la suscripción al evento en cuestión. La propiedad publicación/suscripción del MOM lo hacen idóneo para el control de errores y otro tipo de mecanismos análogos, como los sistemas de alarmas.

Si en cambio es el middleware el que no puede inicializarse (fase 1), debe informarse de inmediato al administrador del sistema o experto tecnológico mediante una alerta de algún tipo, ya que en este caso se trata de un error grave y no puede solventarse por medios convencionales.

A continuación, se comentan brevemente cada una de las etapas correspondientes a este arranque inicial del sistema, presentando por último la Figura 4.7 que ilustra todo el proceso.

1. **Activación del MOM:** se crean las estructuras de datos principales en memoria principal para el manejo de datos, así como las interfaces que sirven de enlace para la comunicación con otros participantes que ya están trabajando en el middleware. En este paso se crea también un evento para el control de errores. De esta forma, los diferentes módulos que van activándose e inicializándose en sucesivas etapas pueden informar de posibles contingencias surgidas mediante la publicación en este evento. Así se consigue centralizar todos los problemas en un determinado espacio, facilitando la labor al experto tecnológico, que puede estar advertido de cualquier circunstancia que se manifiesta en el sistema únicamente con la suscripción a este evento.
2. **Activación del SAB:** se activan todos los módulos constituyentes del bloque de administración del sistema a excepción del módulo de gestión de arranque, que ya se encuentra inicializado. Cada módulo lleva asociado su fichero de configuración correspondiente, que es especificado por el experto tecnológico o administrador de sistema.
3. **Activación de las ORDB y trasvase ORDB/MOM:** se inicializan las bases de datos pertinentes conforme a lo establecido con el módulo de configuración de las ORDB. A continuación, se realiza un volcado de datos de las ORDB al MOM. El objetivo es la creación de repositorios de información que van a estar accesibles desde el MOM. Ejemplos

de repositorios pueden ser aquellos vinculados con arquetipos, pacientes, modelos matemáticos, servicios, o cualquier otro recurso que pueda suministrarse a los usuarios y pueda serles de utilidad. Estos repositorios se presentan como una medida para la localización de recursos por parte de los usuarios, realizándose de forma sencilla a través de suscripciones. Una vez localizado el recurso del que se está interesado, se accede a él normalmente mediante el acceso a las ORDB (si se trata de un recurso de información).

4. Activación del CGB: que al igual que para la activación del SAB, se encarga de inicializar todos los módulos que componen el Bloque de Gestión del Conocimiento.
5. Activación de interfaces con elementos externos al sistema: se activan las interfaces involucradas en la toma de datos de dispositivos a partir del módulo de monitorización del CGB, así como la inicialización de protocolos, interfaces de comunicación y cualquier elemento implicado en la transmisión de datos con entidades ajenas al sistema.
6. Activación de interfaces gráficas: en esta última etapa se arranca el servidor web que va proveer el portal gráfico que ofrece el acceso a los recursos disponibles en el sistema para todos los usuarios.

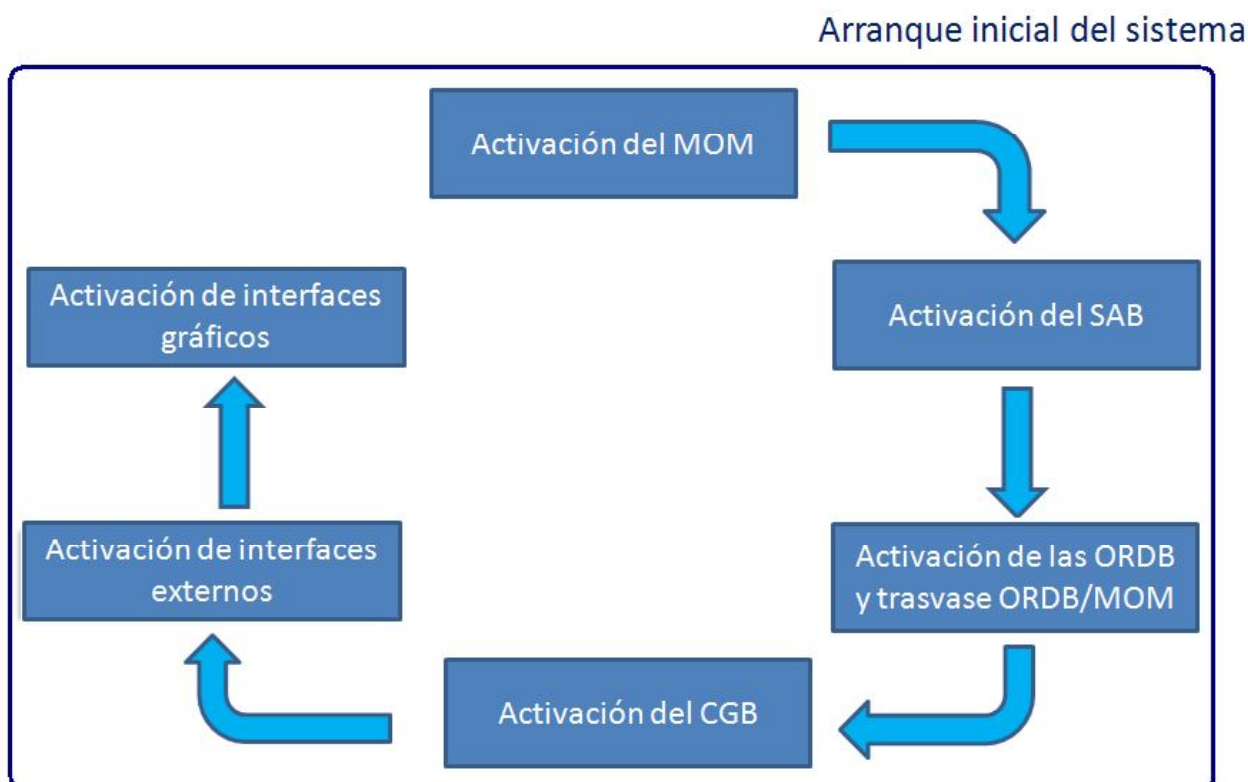


Figura 4.7: Arranque inicial del sistema

Arranque “en caliente” del sistema

Se trata de un conjunto de procesos de este módulo que se pone en marcha justo cuando se completa el proceso de arranque inicial del sistema. Se encarga de detectar, mientras que el sistema está en marcha, posibles cambios en los módulos aledaños pertenecientes al CGB. De esta forma, se va comprobando de forma periódica si existe algún tipo de modificación en los mismos y se gestiona estos cambios de forma segura y de forma transparente para el resto de módulos del sistema.

4.3.2 Módulo de gestión de errores

Uno de los retos que debe abordar este módulo es saber distinguir entre recursos que están fallando y recursos que simplemente están siendo accedidos por muchos usuarios y cuya respuesta puede demorarse. Es importante que posea un mecanismo eficiente de replicación y de balanceo de cargas para evitar estas situaciones, y también definir técnicas que definan un procedimiento a seguir en caso de que se presenten coyunturas de esta índole.

El módulo de gestión de errores está suscrito a un evento dentro del MOM que se encarga de recoger todas las alertas de errores, avisos, informes y demás información de utilidad proveniente de todos los módulos constituyentes del sistema. Debe definirse por tanto una política descriptiva de código de errores, que asocie un determinado número de error con una descripción concreta del mismo.

En ciertos casos estos problemas podrán resolverse de forma autónoma, o en caso contrario, puede decidirse notificar simplemente al experto tecnológico para que él se encargue de solventar el contratiempo en cuestión.

4.3 Middleware Orientado a Mensajería (MOM)

Con el fin de aprovechar las ventajas proporcionadas por los sistemas distribuidos dentro del entorno sanitario, se hace uso del concepto del middleware como centro de conexión entre los diferentes módulos del sistema, e intermediario para el intercambio información entre los mismos. Un middleware es una capa software que ofrece un conjunto de servicios para la comunicación entre aplicaciones distribuidas que funcionan en plataformas heterogéneas. Se encarga de abstraer la complejidad de la redes de comunicación, sistemas operativos y lenguajes de programación, suministrando un API que proporciona transparencia para todas las aplicaciones distribuidas.

En el capítulo anterior se realizó un breve análisis sobre los middleware más utilizados actualmente, con especial atención a los orientados a objetos y a los orientados a mensajería (MOM). Dentro de los MOM, destacaba el paradigma de comunicación publicación/suscripción, que está basado en la tecnología *push*, en donde al contrario que en el modelo cliente/servidor (tecnología *pull*), el publicador es el encargado de iniciar la transacción una vez que el suscriptor se ha inscrito a un determinado canal, mensaje u evento.

El modelo de comunicación publicación/suscripción posee una serie de singularidades que lo hacen idóneo para su utilización en sistemas de alarmas o eventos, que son habituales dentro del ámbito sanitario. Ofrece una potente abstracción para la multidifusión o comunicación en grupo gracias al desacoplo entre nodos productores y consumidores de información, que aportan una mayor flexibilidad en la comunicación mediante el paradigma *many-to-many*. De esta forma, el consumidor de la información no necesita conocer el origen de los datos, ni cómo debe obtenerlos, sino simplemente debe preocuparse por asociarse al canal o evento contenedor de la información que está interesado.

La ventaja anterior también resulta útil para el intercambio de información entre modelos que están ejecutándose al mismo tiempo dentro de un entorno distribuido. La generación y obtención de información de estos modelos pueden aprovecharse del paradigma publicación/suscriptor, que resulta óptimo para este tipo de escenarios.

Igualmente, se consigue también evitar las fases de establecimiento y liberación de conexión presentes en el modelo de comunicación cliente/servidor, omitiendo además la asignación de roles para cada uno de los participantes involucrados. En un MOM, cualquiera puede ejercer de publicador, de consumidor o de ambos al mismo tiempo, por lo que existe ecuanimidad entre todos los nodos, desapareciendo por tanto la dependencia de servidores intermediarios y el concepto de “petición” de servicio. Esta ventaja implica una disminución de la latencia y una mayor eficiencia en el consumo de ancho de banda, virtud que puede beneficiar a dispositivos de monitorización con limitaciones hardware que operan en tiempo real gracias al ahorro de energía que esto conlleva.

Existen diferentes tecnologías que implementan el sistema de distribución MOM. Dentro de la arquitectura propuesta, se deben considerar los siguientes aspectos a la hora de la elección de uno adecuado:

- Se debe de utilizar una tecnología abierta que esté construida bajo una especificación que sea pública y esté ampliamente probada y extendida.
- Se debe proveer un servicio orientado a conexión y otro no orientado a conexión para la transmisión de información confiable y no confiable.
- Se debe garantizar la entrega en orden de los datos. Esto es especialmente relevante cuando se pretende enviar bloques de datos indivisibles, como ficheros de audio, imágenes o vídeo.
- Se debe proporcionar un mecanismo que facilite el trasvase de información del MOM a la base de datos-relacional y viceversa. Con esto se garantiza la persistencia de los datos en memoria secundaria.
- Se debe controlar el acceso a la información de suscriptores *late-joining* (aquellos que se suscriben tarde a un determinado canal u evento), y que están interesados en instancias pasadas que ya han sido publicadas.
- Es recomendable que se ofrezca soporte para operar en tiempo real, permitiendo que tanto publicadores como suscriptores lleguen a un compromiso respecto a unas restricciones temporales.
- Es recomendable que se proporcione un mecanismo de filtrado de la información dentro de un determinado canal u evento, para que los consumidores puedan realizar suscripciones basadas en contenido y obtengan únicamente los datos a los que están interesados.
- Es recomendable que se proporcione un mecanismo de control de flujo para que los publicadores no “ahoguen” a suscriptores con limitada velocidad para aceptar datos, ya sea por características hardware o de red.
- Es recomendable que se defina una política de seguridad que evite la publicación anónima de información en los canales o eventos, así como el uso de un algoritmo para la encriptación de los datos. Si esto no se suministra, debe implementarse en capas superiores.
- Es recomendable la existencia de un protocolo de interconexión que facilite la distribución de información entre diferentes MOM. Con esto se permite la compartición de recursos entre diferentes sistemas, dotando a todo el conjunto de la característica deseable de escalabilidad.
- Es recomendable que se provean pasarelas que favorezcan la adición de nuevas tecnologías que complementen al MOM.
- Es recomendable que se permita la construcción del modelo de datos a utilizar, evitando así el uso de contenedores rígidos de información como mensajes o paquetes, que tienen definida una cabecera que a menudo resulta poco útil.
- Es recomendable que se ofrezca un amplio soporte para un gran número de sistemas operativos, lenguajes de programación y protocolos de comunicación, así como de interfaces normalizados.

4.4 Interfaz de usuario

La gran variabilidad de usuarios que pueden actuar dentro del sistema, y los sistemas heterogéneos involucrados en este proceso, influyen decisivamente en el desarrollo de una interfaz normalizada que abarque todo tipo de exigencias hardware, físicas y las demandables por el usuario. Esta diversidad se ve reflejada en el uso de sistemas operativos dispares, lenguajes de programación soportados, arquitecturas de computador distintas, dispositivos con limitaciones hardware, y usuarios con distintos tipos de necesidades. La interoperatividad llegados a este punto se torna crítica si quieren superarse todos estos contratiempos.

De esta forma, y concretando para el personal sanitario, se debe presentar la información de la manera más intuitiva posible, distinguiendo la información que pueda resultar de interés para los mismos: últimos resultados de laboratorio, acceso a historias clínicas, disponibilidad de documentación, artículos de conocimiento clínico, variables medibles en tiempo real tomadas por dispositivos de monitorización,...

Teniendo en cuenta lo anterior, se ha optado por proveer una GUI (Interfaz Gráfico de Usuario) basado en la Web, debido al auge que ha alcanzado en los últimos años, y a su masificación dentro de todo tipo de sistemas de información. Este éxito ha sido posible (entre otros muchos factores), gracias a la labor de instituciones de estandarización tales como el W3C (World Wide Web Consortium).

Para el acceso a este interfaz Web, se hace indispensable de un sistema de control de acceso basado en credenciales de autorización, y de una política de permisos que debe estar definida dentro del sistema y almacenada dentro de una base de datos de manera consistente en del módulo destinado para dicho fin. Asimismo, los datos relacionados con el usuario, tanto personales como los empleados en su práctica clínica diaria, deben encontrarse encriptados, para garantizar la confidencialidad y privacidad de los mismos.

Básicamente, el proceso de conexión a la interfaz está constituido por los siguientes pasos:

- Autorización de usuario mediante un proceso de logeo: con un nombre de usuario que lleva asociado un identificador único y una contraseña, cifrada mediante algún algoritmo de encriptación de datos.
- Negociación del cifrado de datos: para garantizar que los datos que se transmitan permanezcan seguros e íntegros.
- Disposición de funcionalidades: basado en unas credenciales de autorización, el sistema debe ser capaz de proveer una serie de servicios en función del nivel de acceso del usuario.
- Extracción del perfil de usuario: que extrae las preferencias y las opciones de configuración de usuario. Este perfil de usuario debe seguir un proceso de auto-aprendizaje, adaptándose a las exigencias demandables, y evolucionando en función de la experiencia proporcionada por el usuario. También debe de ofrecerse un mecanismo que posibilite la recuperación de sesiones de trabajo pasadas, permitiendo reutilizar el trabajo anterior y presentando al usuario toda la información de la misma manera a como él mismo la dejó cuando abandonó la interfaz por última vez.
- Restauración de la última sesión
- Inicio de sesión.

En la Figura 4.8 se muestra una ilustración para una mejor comprensión de lo comentado con anterioridad:

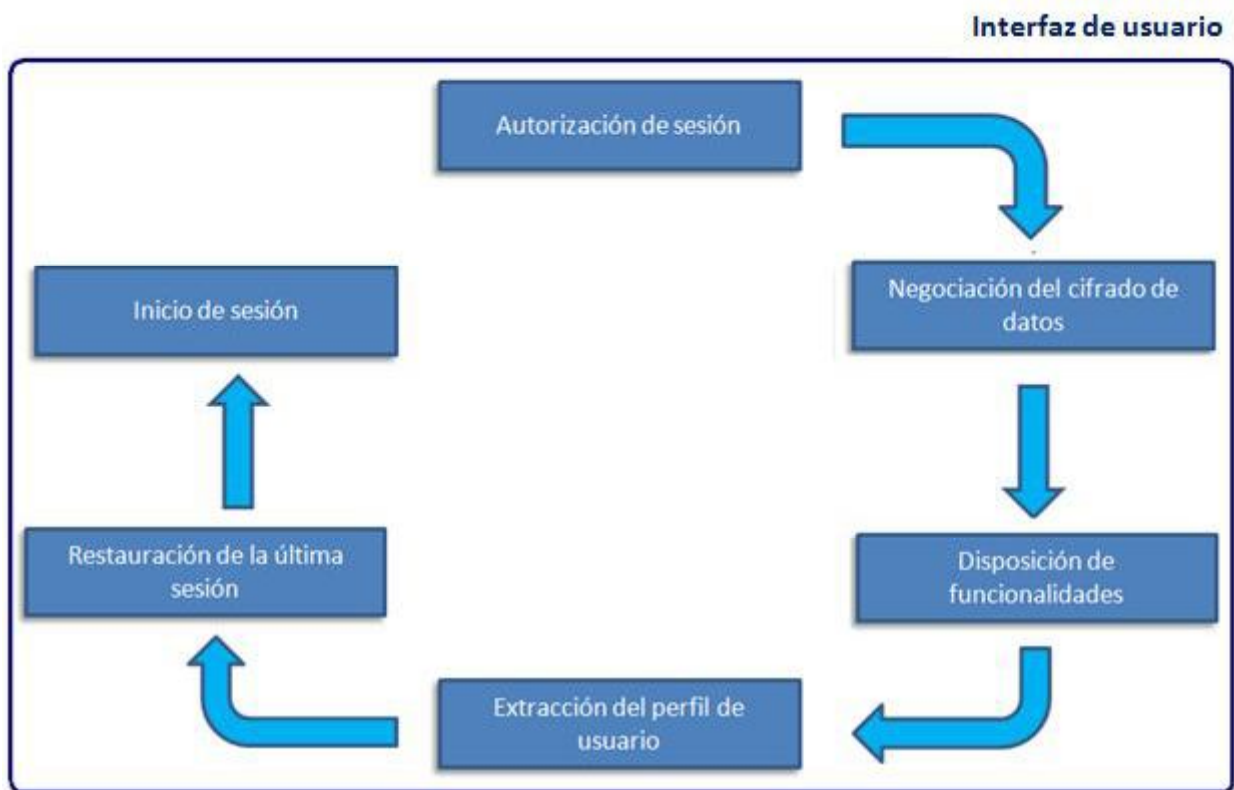


Figura 4.8: Interfaz de usuario