

# INTEGRACIÓN DE REGISTROS CLÍNICOS ELECTRÓNICOS DISTRIBUIDOS POR MEDIO DE ARQUETIPOS DE EXTRACTOS DE HISTORIAS CLÍNICAS Y XML

*Pere Crespo (pedcremo@inf.upv.es), Jose Alberto Maldonado (jamaldo@upvnet.upv.es),  
Montserrat Robles (mrobles@fis.upv.es)*

*Grupo de Informática Médica, E.U.I.  
Universidad Politécnica de Valencia*

## RESUMEN

Cada día más, la responsabilidad de la atención sanitaria de un paciente está compartida por un conjunto de profesionales sanitarios pertenecientes a diversas disciplinas e instituciones, por tanto, es vital el poder compartir la información sanitaria sobre los pacientes. Pero, actualmente esta información está repartida en múltiples sistemas de información heterogéneos y autónomos, por tanto, el acceso uniforme a los registros clínicos distribuidos es una tarea problemática. Este trabajo tiene como objetivo describir un sistema de integración de registros clínicos electrónicos desarrollado por el grupo de informática médica del BET de la Universidad Politécnica de Valencia.

La aproximación tecnológica utilizada para el diseño e implementación del sistema de integración de registros clínicos electrónicos desarrollado se basa en el estándar europeo de arquitectura de historia clínica electrónica ENV13606 del CEN/TC251. El uso de este estándar en un proyecto de integración se debe a que consideramos que el desarrollo de sistemas de historias clínicas electrónicas se prevé a medio plazo y por tanto cualquier proyecto de integración debe tener en cuenta sus implicaciones e intentar ser compatible con ellas. Asimismo, la estandarización de los registros clínicos, tanto a nivel de estructura como de contenido y significado es vital si los datos clínicos deben ser enviados fuera del departamento u organización donde fueron creados, de forma que el destinatario puede interpretar fielmente el significado original de los datos.

El estándar europeo se utiliza como modelo de referencia para la definición de arquetipos o plantillas de extractos válidos de historias clínicas electrónicas. Estos arquetipos, que son definidos por los propios usuarios del sistema en función de sus necesidades, permiten hacer pública la información almacenada en las distintas bases de datos y a la vez dotarla de un significado clínico, por tanto, la verdadera integración se realiza al nivel de metainformación (arquetipo) en vez de al nivel de datos. La representación y compartición de la información se realiza a través de XML, el cual se ha convertido en el estándar de facto para el intercambio de información entre aplicaciones a través de Internet.

## INTRODUCCIÓN

El trabajo llevado a cabo por el grupo ha consistido en el desarrollo de un sistema informático cuyo propósito es facilitar el acceso a los profesionales sanitarios de una forma integrada a toda la información dispersa por los diferentes Sistemas de Información (SI) departamentales de un hospital.

La aproximación tecnológica utilizada para solucionar el problema ha sido la de desarrollar *un repositorio lógico*, el cual se encarga de extraer la información requerida de los S.I conectados en un momento dado. Por repositorio lógico de datos clínicos entendemos un sistema capaz de integrar toda la información clínica sobre un paciente dispersa en múltiples lugares de forma que los usuarios finales del sistema tengan la impresión que están trabajando con una única base de datos que contiene toda la información (historia clínica), lo denominamos lógico porque esta base de datos no existe físicamente. Por el contrario, en un repositorio físico si que existe una base de datos física (o un conjunto de bases de datos) que contiene toda la información y los usuarios finales realizan consultas a esta base de datos.

Las ventajas de un repositorio lógico frente a uno físico en un proyecto de integración de datos clínicos, no resultan nada desdeñables:

- ? Se evita la existencia de datos redundantes, cuya existencia puede acarrear problemas a la hora de asegurar la coherencia entre las diversas instancias de un mismo dato.
- ? E es una solución más genérica sin necesidad de una versión AD-HOC para cada hospital
- ? Es una solución escalable a la incorporación de nuevos sistemas departamentales.
- ? Al haber una capa lógica por encima de los sistemas, se puede construir y presentar bajo la arquitectura que se desee. En nuestro caso lo hacemos bajo la arquitectura basada en el estándar ENV13606 ya que confiamos que algún día llegará a ser un estándar europeo de Historia Clínica Electrónica.
- ? El desarrollo de un repositorio lógico facilita en gran manera el desarrollo de repositorios físicos de datos clínicos (persistencia de los datos extraídos).

La solución adoptada para la integración de información clínica distribuida se ha basado en respetar la autonomía de los sistemas de información ya existentes (cuanto menos cambios mejor), que sea fácilmente escalable y tenga en consideración los estándares existentes. El sistema desarrollo, básicamente es una capa de software que se sitúa por encima de los sistemas departamentales ya existentes y permite gestionar una historia clínica electrónica “virtual”. Es virtual porque los usuarios finales pueden disponer de la HCE del paciente pero realmente no existe como entidad propia en ningún sistema informático sino que se encuentra repartido por múltiples bases de datos departamentales.

## **BREVE INTRODUCCIÓN AL PRESTÁNDAR ENV13606**

La arquitectura definida por el estándar da las reglas necesarias para construir cada una de las piezas de las que se compone un arquetipo. Cada una de estas piezas que da cuerpo a un documento o agregado de información las podemos clasificar de la siguiente manera: EHCR, carpeta, composición, encabezado, clúster e ítem de información. A continuación presentamos una breve descripción de cada una de ellas, solo se definirán aquellas componentes relevantes e instanciables dentro del conjunto del estándar, hacemos referencia al [1] para la obtención de mayor información en cuanto a la arquitectura de la historia clínica electrónica (HCE) se refiere.

**EHCR, Root Architectural Component:** Representa la raíz de la historia clínica, es decir la carpeta general donde se podrá encontrar toda la información referente a un mismo paciente. De ella dependerán el resto de componentes que formarán la HCE.

A partir de dicha carpeta se agruparán todas las componentes necesarias para dar forma a la HCE, es decir, el EHCR es el nodo principal de una estructura en árbol de la que dependerán o colgarán el resto de componentes.

Se han implementado dos clases diferenciadas de componentes, una cuyo fin es agrupar otras componentes que cuenten con información de contexto similar o que poseen características similares y otro grupo de componentes que representa a la unidad de información que es indivisible dentro de un mismo contexto, es decir representa la unidad estructural mas pequeña con significado en la que se puede descomponer una HCE.

Dentro las clases de componentes que se encargan de la estructura y agrupamiento de los ítems de información se encuentran:

- *Folder* (Carpeta), Se usa como subdivisión de alto nivel de la HCE de un paciente. Agrupa entradas referidas a un periodo de tiempo completo, un mismo departamento o problema de salud en particular. Ejemplos de esta componente sería una estancia hospitalaria o una historia de atención primaria.

- ? Composition (Composición), Contiene un nivel más homogéneo de componentes, que poseen en común una fecha, un lugar de atención sanitaria o una sesión. Representa la idea de un documento. Ejemplos de una composición pueden ser una hoja de anestesia, un informe de quirófano o un informe de alta.
- ? Headed Section (Sección con encabezamiento), representa subdivisiones dentro de las composiciones, cada subdivisión contendrá un tema en común o un mismo proceso sanitario. Ejemplos, Plan de tratamientos, dietas prescritas o síntomas. Pueden verse como subdivisiones de un documento.
- ? Clúster, representa conceptos compuestos indivisibles, formados por ítems u otros clústers. Ejemplos de esto son la presión sanguínea (sístole, diástole) o una receta compuesta por dos medicamentos.

Clases de ítems de información o unidades mínimas con significado dentro de la HCE:

- ? Item de texto: Representa un texto que proporciona una descripción sobre un contenido. Dicho contenido puede formar parte de un componente complejo original.
- ? Item de dirección: La dirección del paciente o de un grupo de pacientes relacionados.
- ? Item de identificador de persona: Información demográfica sobre una persona. Incluye campos como nombre, fecha de nacimiento, sexo, o dirección.
- ? Item con códigos: Un ítem de un hospital u otra entidad sanitaria relacionada representada mediante un código o composición de códigos con texto adicional opcional.
- ? Item de medicación: información sobre un tratamiento previo, planificado o actual con una sustancia o compuesto químico para la prevención o tratamiento de una condición médica.
- ? Item de observación cuantificable: representa el resultado de una medida clínica o de un laboratorio de investigación. Se definen cuatro tipos distintos de estos:
  - ? Rango numérico: La medida viene expresada en un rango entre dos valores numéricos calificados por una unidad de cantidad.
  - ? Valor numérico: Una medida expresada como un valor numérico y una unidad de cantidad con un comparador matemático opcional.
- ? Texto: Un resultado de una observación o investigación expresado usando texto o códigos.
- ? Fecha: Un resultado de una observación o investigación expresado como una fecha. Ejemplo: Fecha estimada de entrega de una obstetricia.

## ARQUETIPOS

Formalmente un arquetipo es una definición basada en un modelo de referencia de una estructura de información utilizada en un dominio particular. Los profesionales sanitarios suelen manejar un conjunto más o menos fijo de estructuras de información que representan conceptos médicos para la realización de sus actividades, por ejemplo, informe de alta, historia clínica de primaria, resultados bioquímicos, diagnóstico, etc. Estos conceptos médicos, pueden definirse formalmente por medio de arquetipos si tomamos como modelo de referencia por ejemplo una AHCI. Obviamente, la definición de arquetipos es tarea de los especialistas en el campo de interés, así por ejemplo, los patólogos puede definir arquetipos para la representación de resultados bioquímicos.

Los arquetipos son especificaciones formales que pueden ser utilizados por los sistemas de información para crear estructuras de datos y validar la entrada de datos de forma que se puede garantizar no sólo que los datos introducidos cumplen el modelo de referencia sino también las restricciones definidas por el arquetipo. Una consecuencia muy interesante es la posibilidad de poder desarrollar sistemas de información basados en el modelo de referencia, y donde los conceptos del dominio, que están modelados como arquetipos, son definidos y usados por el sistema en tiempo de ejecución, por tanto, es mucho más difícil que el sistema quede obsoleto. Dos de las consecuencias más interesantes de la utilización de arquetipos son: su utilización para compartir información entre sistemas permite una comunicación a nivel de conocimiento en vez de simplemente a nivel de datos y posibilita el desarrollo de potentes herramientas para la consulta de información.

## SISTEMA DE INTEGRACIÓN

Los arquetipos forman el núcleo básico de la nuestra solución para la integración de sistemas de información hospitalarios. Su propósito es hacer público la información clínica almacenada en las bases de datos subyacentes, y al mismo tiempo ocultar su detalles técnicos (heterogeneidad). Por tanto, forman una capa semántica sobre las bases de datos a integrar y permiten asociarles una semántica específica. Por tanto la verdadera integración se consigue al nivel de metainformación en vez de al nivel de datos. Una de las principales ventajas de esta aproximación es que cada hospital o servicio puede definirse su propia vista de la historia clínica.

Ya que información clínica sobre los pacientes reside en las bases de datos integradas, se debe definir un mecanismo que permita relacionar los arquetipos con los esquemas de las bases de datos. Es importante resaltar que es no sólo necesario recuperar la información puramente clínica pura también cualquier información de contexto disponible con el propósito de salvaguardar el significado original de datos. ENV13606 describe que la información del contexto debe acompañar cualquier extracto de historia clínica. Para facilitar el intercambio de información clínica, los datos relacionales deben transformarse en XML. Esta transformación no es una tarea trivial, ya que los dos modelos difieren significativamente. Los datos relacionales son planos, están normalizado en muchas relaciones, y su esquema es a menudo propietario. Por el contrario, los la información en XML esta jerarquizada, no está normalizada, y su esquema es público, en nuestro caso ENV13606. Esta transformación implica el concatenar tablas (operación join del álgebra relacional), seleccionar y proyectar los datos que deben ser exportados, hacer corresponder los nombres de los campos y tablas con elementos y atributos de XML, crear jerarquías de elementos y posiblemente procesar algunos valores.

Utilizamos el concepto de vistas como herramienta para convertir los datos relacionales en extractos de historias clínicas (documentos XML que cumplen el estándar). Una vista en nuestro sistema, esta formada por un conjunto de correspondencias de valores entre los atributos de los arquetipos (elementos o atributos XML) y los campos de las bases de datos, una o varias consultas en SQL y un conjunto de campos pivote. Los diseñadores de arquetipos deben define un conjunto de estas correspondencias para definir completamente un arquetipo. A partir de este conjunto de correspondencias el sistema es capaz de generar una consulta [2]candidata que permitan la instanciación de los arquetipos con datos clínicos para un paciente y determinar qué campos sirven como identificador único de un arquetipo o cualquier atributo multivaluado, estos campos son los que denominamos campos pivotes. Los campos pivotes son básicos para poder generar XML eficientemente ya que nos permiten determinar cuando nos hemos encontrado con una nueva instancia de un elemento.

Al generar las consultas candidatas para un arquetipo se han tenido en cuenta los siguientes aspectos:

- Estamos intentando generar extractos de historias clínicas por tanto no toda la información contenida en las bases de datos es relevante, más concretamente, estamos únicamente interesados en información para la cual podamos determinar qué paciente se trata.
- Las relaciones existentes entre los datos, expresados típicamente por claves ajenas, deben ser mantenidas
- No debemos perder la información.
- La pregunta debe facilitar la posterior generación de documentos XML.

## ARQUITECTURA DEL SISTEMA

El sistema informático desarrollado es complejo entramado de servidores y clientes cada uno de ellos diseñado para desempeñar una función específica, véase figura 1. El sistema se sitúa entre los usuarios y las bases de datos a integrar, por tanto se puede ver como una forma de middleware, el cual recupera, bajo demanda de los usuarios, la información relevante sobre el paciente y la entrega a los usuarios. El sistema genera un documento XML que contiene la información clínica demandada y que está estructurada según el estándar europeo de arquitectura de historia clínica. El servidor de HCE presenta una interface sencilla que puede ser invocada por medio de SOAP, esto dota de gran independencia a la hora de desarrollar aplicaciones clientes (por ejemplo estaciones clínicas). Las aplicaciones clientes son las que procesan y presentan a los usuarios finales los resultados según sus necesidades, por ejemplo pueden transformar fácilmente la información en HTML, PDF, RTF, DOC etc.

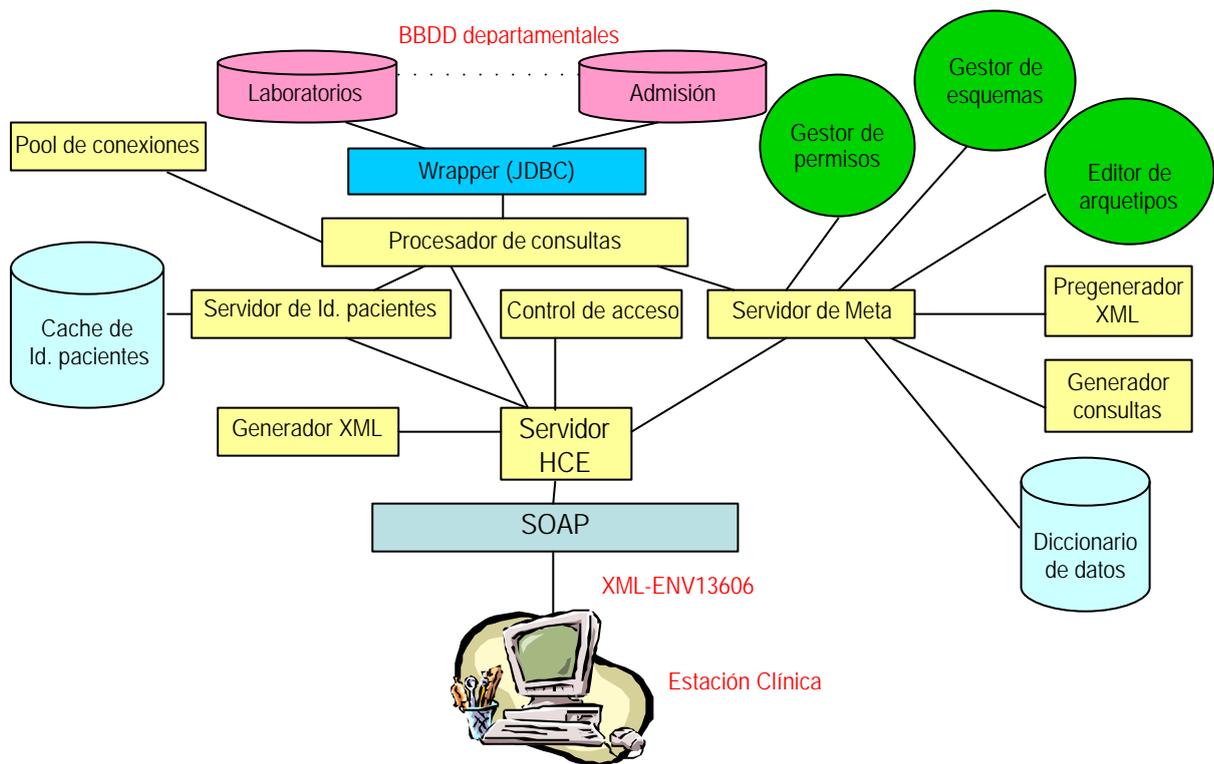


Figura 1. Arquitectura del sistema de integración desarrollado

Una de las componentes principales del sistema es el servidor de metainformación, el cual gestiona el diccionario de datos donde se almacena toda la información necesaria para el funcionamiento del sistema. El diccionario de datos contiene información sobre las bases de datos conectadas al servidor, los esquemas de las bases de datos, qué objetos de las bases de datos conectadas pueden ser accedidos por las aplicaciones clientes, la definición y versiones anteriores de los arquetipos, los enlaces de éstas con los esquemas de bases de datos, las relaciones entre arquetipos y sinónimos.

El servidor de historias médicas puede utilizar otros dos servidores para realizar su tarea, el servidor de control de acceso previene que usuarios no autorizados puedan acceder a los datos y el servidor de identificadores de pacientes que permite identificar pacientes idénticos entre diferentes sistemas cuando no existe un identificador de paciente universal. Se han desarrollado, también, dos aplicaciones visuales para la gestión de la metainformación del sistema: el editor de arquetipos y el gestor de esquemas de bases de datos. El primero tiene como propósito el ayudar en el proceso de diseño y validación de arquetipos. El segundo es una herramienta para gestionar y completar los esquemas de las bases de datos conectadas al servidor.

## CONCLUSIONES

En este artículo hemos hablado sobre la integración de sistemas de información hospitalarios haciendo uso del estándar ENV13606. Consideramos que el desarrollo de sistemas de historias clínicas electrónicas se prevé a medio plazo y por tanto cualquier proyecto de integración debe tener en cuenta sus implicaciones e intentar ser compatible con ellas. Asimismo, la estandarización de los registros clínicos, tanto al nivel de estructura como de contenido y significado es vital si los datos clínicos deben ser enviados fuera del departamento u organización donde fueron creados, de forma que el destinatario pueda interpretar fielmente el significado original de los datos.

El sistema es totalmente escalable, cualquier incorporación o modificación de un S.I departamental en un hospital podrá ser incorporado al sistema sin ningún problema.

Es un sistema multiplataforma que puede funcionar prácticamente bajo cualquier arquitectura o S.O porque ha sido implementado con el lenguaje de programación JAVA

Queda totalmente abierta la implementación de las “Estaciones Clínicas” para que implementen en el lenguaje que se desee, siempre que este lenguaje tenga las herramientas necesarias para utilizar SOAP y XML.

Aunque el sistema, en esta primera fase, ha sido concebido para integrar y servir información dentro de un único hospital, consideramos que la solución adoptada puede ser fácilmente ampliada para permitir la compartición de información clínica entre diversas instituciones.

## BIBLIOGRAFIA

- [1] CEN/TC251 WG I.: Health Informatics-Electronic Healthcare Record Communication- Parts 1, 2, 3 y 4. Final Draft prENV13606, 1999.
- [2] Rajaraman, J.D. Ullman, “Integrating information by outerjoins and full disjunction”, PODS, pp. 238-248, 1996.