

**SISTEMA DE PROCESAMIENTO TRANSACCIONAL Y
ANALÍTICO PARA LA VIGILANCIA EPIDEMIOLÓGICA
DEL DENGUE**

*Tesis en opción al título académico de Máster en
Matemática Aplicada e Informática para la Administración*

Eduardo López Hung



UNIVERSIDAD DE HOLGUÍN
Facultad de Informática y Matemática

**SISTEMA DE PROCESAMIENTO TRANSACCIONAL Y ANALÍTICO
PARA LA VIGILANCIA EPIDEMIOLÓGICA DEL DENGUE**

*Tesis en opción al título académico de
Máster en Matemática Aplicada e Informática para la Administración*

Autor: Ing. Eduardo López Hung

Tutores: Ph. D. Raúl Ernesto Menéndez Mora
Dr. C. Mauro Misael García Pupo

Holguín. 2012

*“La tecnología nos ofrece vastas posibilidades,
pero sólo si tenemos la imaginación suficiente
para aprovecharlas”.*

Arno Penzias

DEDICATORIA

A mi esposa, quien se regocija por resultados como estos.

A mis padres, mi hermano y familiares, quienes siempre han estado al tanto de mi formación y superación profesional.

AGRADECIMIENTOS

A Dios por su guía y fortaleza.

Al claustro de la maestría, especialmente a mis tutores Raúl y Mauro, y a la incansable y siempre querida profesora Rosa I. Urquiza Salgado por sus conocimientos, consejos y ayuda en todo momento.

A mi esposa que siempre estuvo a mi lado ayudándome y apoyándome en todo cuanto me hizo falta.

A los compañeros del Centro Provincial de Higiene y Epidemiología de Santiago de Cuba por su ayuda y participación en el desarrollo de la tesis.

A mi familia, amigos, y compañeros de trabajo que estuvieron siempre al tanto de mi formación en este programa de maestría.

A todos muchas gracias.

RESUMEN

En una era de amenazas por enfermedades emergentes y reemergentes, la administración en Salud Pública presta especial atención a la situación epidemiológica de las mismas. En Santiago de Cuba, el Centro Provincial de Higiene y Epidemiología es el encargado de controlar dicha situación, donde la gestión de la información subyacente en el gran volumen de datos generado, presenta deficiencias en detrimento de las decisiones administrativas que pudieran ser tomadas. Por lo que se planteó como problema científico la no existencia de una vía que garantice de forma dinámica el procesamiento, evaluación y presentación de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, que tribute al proceso de toma de decisiones administrativas en el orden epidemiológico.

En consecuencia se propuso como objetivo: elaborar un sistema informático que permita el procesamiento transaccional y analítico en línea de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, que tribute al proceso de toma de decisiones administrativas en el orden epidemiológico; asumiendo la conveniencia del desarrollo del mismo utilizando datawarehouse y análisis OLAP, proponiendo el Pentaho Business como plataforma de desarrollo, y MySQL como Sistema Gestor de Bases de Datos.

Se elaboró un sistema informático que apoya el proceso de toma de decisiones en el orden epidemiológico, caracterizado por su flexibilidad, capacidad de análisis, compatibilidad y robustez, enfocado a la explotación de grandes volúmenes de datos del Sistema de Vigilancia Epidemiológica del Dengue.

ABSTRACT

In an era of threats from emerging diseases, public health administration pays special attention to the epidemiological situation generated by them. In Santiago de Cuba, the Provincial Center of Hygiene and Epidemiology is responsible for controlling the situation, where the underlying information management from the large volume of data produced, is deficient at the expense of administrative decisions that may be made. As stated as a scientific problem nonexistence of a way to ensure a dynamic processing, evaluation and information related to the Epidemiological Surveillance System of Dengue in Santiago de Cuba, that tribute to the decisions – making process from an epidemiological order.

Consequently the purpose of this research was: to develop a computer system for online transactional and analytical processing, related to the Epidemiological Surveillance System of Dengue in Santiago de Cuba, that tribute to the decisions – making process, with the advisability of assuming the development of the system by using data warehouse and OLAP analysis techniques, considering the Pentaho Business as the development platform, and MySQL as the System Database Manager.

It was developed a computer system, that supports decision – making process from an epidemiological order, characterized by its flexibility, analysis capability, support and strength; focused on the exploitation of a large set of data related to the Epidemiological Surveillance System of Dengue.

INDICE

INTRODUCCIÓN.....	1
CAPÍTULO 1. FUNDAMENTOS TEÓRICOS DEL SISTEMA DE VIGILANCIA EPIDEMIOLÓGICA DEL DENGUE Y HERRAMIENTAS PARA LA AUTOMATIZACIÓN DE SU GESTIÓN.....	9
1.1. El Sistema de Vigilancia Epidemiológica del Dengue	9
1.2. Automatización del Sistema de Vigilancia Epidemiológica del Dengue	11
1.3. Sistemas de apoyo a la toma de decisiones	14
1.4. Inteligencia de negocio para el apoyo a la toma de decisiones	15
1.4.1. Almacén de datos o datawarehouses.....	18
1.4.2. Elementos de un modelo multidimensional	20
1.4.3. Procesamiento Analítico en Línea	22
1.5. Tendencias y tecnologías actuales de desarrollo.....	23
1.5.1. Modelo Cliente – Servidor.....	23
1.5.2. Tecnologías de inteligencia de negocio de código abierto	25
1.5.3. Pentaho Open Source Business Intelligence.....	26
1.5.4. Proceso Unificado de Desarrollo de Software	29
1.6. Conclusiones del capítulo.....	31
CAPÍTULO 2. CONSTRUCCIÓN DEL SISTEMA DE PROCESAMIENTO TRANSACCIONAL Y ANALÍTICO PARA LA VIGILANCIA EPIDEMIOLÓGICA DEL DENGUE	32
2.1. Breve caracterización del negocio desde sus procesos	32
2.2. Requerimientos del sistema	33
2.2.1. Requerimientos funcionales y no funcionales.....	33
2.2.2. Actores del sistema.....	39
2.2.3. Diagrama de paquetes y casos de uso del sistema	40
2.3. Valoración de sostenibilidad de la solución propuesta	42
2.4. Construcción del sistema.....	44
2.4.1. Diagrama de clases	46
2.4.2. Diagrama de clases persistentes y modelos de datos.....	47
2.4.3. Diseño visual.....	49
2.4.4. Soporte para el análisis OLAP en el sistema.....	50
2.4.5. Tratamiento de errores y seguridad.....	55
2.4.6. Despliegue del sistema.....	56

2.4.7. Estructura general del sistema	57
2.5. Utilización del sistema para el apoyo a la toma de decisiones.....	59
2.6. Valoración del sistema utilizando criterio de expertos.....	61
2.7. Conclusiones del capítulo.....	63
CONCLUSIONES.....	65
RECOMENDACIONES	66
REFERENCIAS BIBLIOGRÁFICAS	67
BIBLIOGRAFÍA.....	70
ANEXOS.....	I

LISTADO DE ANEXOS

ANEXO 1. Reglas del negocio.....	I
ANEXO 2. Diagrama de actividades para la vigilancia higiénico – epidemiológica.....	III
ANEXO 3. Diagramas de casos de uso del sistema	IV
ANEXO 4. Diagramas de clases.....	VI
ANEXO 5. Modelos de datos.....	IX
ANEXO 6. Pantallas típicas del sistema.....	XI
ANEXO 7. Esquema Mondrian (XML) donde se definen los cubos creados.....	XIII
ANEXO 8. Encuesta para determinar el coeficiente de competencia de los expertos	XV
ANEXO 9. Encuesta aplicada a los expertos	XVI
ANEXO 10. Resultados generales de la encuesta aplicada a los expertos	XVIII

LISTADO DE FIGURAS

Figura 1.1. Proceso de Inteligencia de Negocio.	17
Figura 2.1. Diagrama de casos de uso del negocio	33
Figura 2.2. Actores del sistema	40
Figura 2.3. Diagrama de paquetes del sistema	41
Figura 2.4. Diagrama de clases persistentes del subsistema OLTP y Administración.....	47
Figura 2.5. Diagrama de clases persistentes para el subsistema OLAP.....	49
Figura 2.6. Diagrama que define el job <i>actualizar_dw_dengue</i>	51
Figura 2.7. Diagrama que define la transformación <i>actualizar_dimensiones</i>	51
Figura 2.8. Diseño visual de uno de los cubos en el Schema Workbench.....	52
Figura 2.9. Vista de análisis de uno de los cubos creados y publicados.....	54
Figura 2.10. Muestra de uno de los reportes creados, publicados y generados	54
Figura 2.11. Diagrama de despliegue del sistema	56

LISTADO DE TABLAS

Tabla I. Actores del negocio	33
Tabla II. Actores del sistema	39
Tabla III. Módulos del sistema	40



INTRODUCCIÓN

En los últimos años ha tenido lugar en el mundo la emergencia o reemergencia de numerosos eventos epidemiológicos. Entre estos eventos se encuentra el descubrimiento de nuevas enfermedades infecciosas, sus agentes etiológicos y su fisiopatogenia, así como otras enfermedades que tuvieron determinados niveles de control y ahora se muestran con incidencias cada vez más altas, convirtiéndose en problemas sanitarios de primera magnitud, tanto en los países en vías de desarrollo como en los desarrollados [Suárez, 2000].

En 1992, el Instituto de Medicina de los Estados Unidos definió como enfermedades emergentes aquellas cuya incidencia se ha incrementado desde las últimas dos décadas, o amenaza con incrementarse en el futuro [Suárez, 2000].

Las enfermedades reemergentes, por su parte, se refieren al resurgimiento de algunas que habían sido aparentemente erradicadas o su incidencia disminuida. Son todas aquellas enfermedades infecciosas conocidas que, después de no constituir un problema de salud, aparecen a menudo cobrando proporciones epidémicas [Suárez, 2000]. En este grupo, se destaca por su incidencia y actualidad la enfermedad del Dengue, la cual es producida por un virus.

Hasta el momento se han descrito cuatro serotipos de este virus que circulan principalmente en países del sudeste asiático, del Pacífico Occidental, y de América Latina y el Caribe, por lo que la enfermedad se considera tropical.

El debilitamiento de los sistemas de salud pública a nivel mundial, debido a la privatización desmedida y la falta de programas sostenibles para el control del vector, han conllevado a que el dengue se convierta en una enfermedad endémica al sur del Río Bravo, es decir, en una enfermedad habitual; a tal punto que en la actualidad los únicos países latinoamericanos donde el dengue no es endémico son Chile, Cuba y Uruguay [Kouri, 2011].



En Cuba, aunque la enfermedad del dengue (entre muchas otras enfermedades transmisibles) no es endémica; esta se aborda justamente desde una perspectiva más amplia, donde la responsabilidad de los profesionales de la salud no se circunscribe a mantener o devolver la salud de los pacientes o a prevenir su alteración, sino que se ensancha hasta el más vasto terreno de la salud de la población, que es sin duda uno de los bienes nacionales más importantes.

En consecuencia, los diagnósticos no se hacen a este nivel con datos extraídos del examen físico o con cifras provenientes de exámenes complementarios; ni las terapéuticas se evalúan sobre la base de la evolución de un paciente; sino que existen indicadores cuyo estudio y análisis revelan el estado de salud de la población en general.

El primer paso para controlar el dengue es la detección e identificación inmediata. Para ello es esencial contar con un sistema organizado de vigilancia; por lo que se ha trabajado en el perfeccionamiento de los programas de prevención y control. En este sentido estos programas se convierten en un pilar fundamental que garantiza una eficiente y rápida vigilancia, que detecta y señala riesgos, y brinda información a todos los que deben conocerla para una oportuna y adecuada toma de decisiones [Suárez, 2000].

La administración en Salud Pública presta especial atención a esta situación epidemiológica, principalmente con el establecimiento de un fuerte y sólido sistema de vigilancia epidemiológica. Este sistema se apoya para su funcionamiento en la estructura del Sistema Nacional de Salud, y desde el año 1986 se viene perfeccionando a través del desarrollo del sistema de información para la vigilancia epidemiológica [Andrés, 1999]. Además ha fortalecido las técnicas y herramientas de diagnóstico de la enfermedad como soporte fundamental del mismo.

De ahí que revista gran importancia el hecho de contar con métodos específicos, rápidos, tempranos y baratos, que permitan por una parte el diagnóstico etiológico



en el paciente; y por otra brindar un pronóstico del comportamiento epidemiológico de la enfermedad en tiempo y espacio.

Bajo esta premisa, el estricto control y vigilancia epidemiológicos de la enfermedad del dengue permite a los distintos niveles de salud la correcta y oportuna toma de decisiones respecto a la situación epidemiológica provocada por esta enfermedad transmisible; al considerar la salud de la población como un bien de incalculable valor, que puede y debe ser administrado.

Si bien esta gestión administrativa debe desarrollarse adecuadamente, ello depende en gran medida de la calidad de la información acerca del objeto administrable, y de los procedimientos empleados para representar, analizar e interpretar dicha información.

En Santiago de Cuba, el Centro Provincial de Higiene y Epidemiología (CPHE) es la institución encargada de llevar a cabo el registro y control de los casos reportados por el Sistema de Vigilancia Epidemiológica del Dengue. Para esta actividad se reciben en el laboratorio del centro las diferentes muestras de sangre de los pacientes con síntomas relacionados con esta enfermedad. A estas muestras se les realiza una serie de procedimientos (denominados determinaciones), relacionados con el "Sistema Ultra Micro Analítico (SUMA)" para la detección de anticuerpos IgM a los cuatro serotipos o cepas del virus del Dengue en suero. Esta prueba de laboratorio confirma el diagnóstico presuntivo de la enfermedad.

Sin embargo, aun cuando se cuenta con una base biotecnológica sólida e irrefutable para el diagnóstico de esta enfermedad, el acierto en las conclusiones y el éxito de las medidas que se deriven de este sistema depende, tanto de la calidad de la información y de las fuentes, como de las técnicas y procedimientos que se utilicen para su análisis e interpretación. Para ello se tiene en cuenta: la cantidad de datos procesados, cantidad de información generada, periodicidad, perspectivas de análisis; así como la confidencialidad y disponibilidad de la información.

Es aquí donde el proceso de toma de decisiones en el CPHE de Santiago de



Cuba, más que depender de las capacidades cognitivas del personal que labora en él, necesita una herramienta que apoye dicho proceso, debido a la cantidad de variables e indicadores involucrados y a sus interdependencias. En tal sentido los avances tecnológicos muestran poderosas herramientas capaces de automatizar, manipular e imitar procesos cognitivos, que mejoran considerablemente la toma de decisiones en muchas situaciones [Buchanan, 2006]. Además tal aspiración está en consonancia con la política del Ministerio de Salud Pública para el establecimiento y desarrollo de la informatización en este sector.

A partir de las premisas anteriores, dentro del funcionamiento del Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, en detrimento de las decisiones administrativas que pudieran ser tomadas en el orden epidemiológico, se identificaron varias deficiencias, entre las que pueden destacarse las siguientes:

- La actividad de registro y control de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba se realiza de forma semi – automatizada, utilizando un fichero en Microsoft Excel, donde se reflejan datos de cada uno de los casos sometidos a las pruebas de laboratorio mediante la tecnología SUMA. Este es consolidado en otro fichero en Microsoft Excel, el cual es replicado en los departamentos del centro y enviado a la Dirección Provincial y Nacional de Salud Pública. Esto trae como consecuencia descentralización, inconsistencia, desactualización constante y la no integridad de la información, que se evidencia en errores en los flujos de información hacia los diferentes niveles, no correspondencia entre los partes emitidos, así como la pérdida frecuente de datos valiosos respecto a este dominio de aplicación.
- Este Sistema de Vigilancia Epidemiológica del Dengue en el territorio genera un gran volumen de datos, al que no se le da el adecuado tratamiento para su correcto almacenamiento y recuperación, a fin de apoyar el proceso de toma de decisiones epidemiológicas pertinentes



según los elementos e indicadores de análisis; provocando en ocasiones la adopción de políticas y acciones que no están en correspondencia con la situación epidemiológica existente de esta enfermedad.

- Los informes y balances son realizados sin ningún apoyo matemático – computacional, a pesar de contar con un número considerable de sistemas de cómputo, y de las ventajas y beneficios que este acarrearía para el apoyo al proceso de toma de decisiones desde el análisis, interpretación y presentación de la información.

De la situación problemática anterior, se identificó el siguiente **problema científico**: ausencia de una vía que garantice de forma dinámica el procesamiento, evaluación y presentación de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, y que tribute al proceso de toma de decisiones administrativas en el orden epidemiológico.

El mismo se enmarca en el **objeto de estudio**: Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba; identificando como **campo de acción**: informatización del Sistema de Información relacionado con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba.

Para resolver el problema científico identificado, se propuso el siguiente **objetivo**: elaborar un sistema informático que permita el procesamiento transaccional y analítico en línea de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, que tribute al proceso de toma de decisiones administrativas en el orden epidemiológico.

Se plantearon las siguientes **preguntas científicas**:

1. ¿Cuáles son las posiciones teóricas en cuanto a la gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue?
2. ¿Cuál es el estado actual del uso de las herramientas informáticas para la gestión de la información relacionada con el Sistema de Vigilancia



Epidemiológica del Dengue en Santiago de Cuba, y las tendencias para su desarrollo?

3. ¿Cómo valorar la pertinencia y sostenibilidad del sistema informático propuesto?
4. ¿Cómo realizar el procesamiento transaccional y analítico en línea, de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba?
5. ¿Cómo valorar la relevancia del sistema informático implementado?

Para el cumplimiento del objetivo de la investigación, responder a las preguntas científicas y de manera general guiar la investigación, se propusieron las siguientes **tareas científicas**:

1. Elaborar los fundamentos teóricos referentes a la gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue.
2. Diagnosticar el uso de herramientas informáticas para la gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, y las tendencias actuales para su desarrollo.
3. Realizar una valoración de pertinencia y sostenibilidad del sistema informático propuesto.
4. Elaborar el sistema informático de procesamiento transaccional y analítico en línea de la información, para el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba.
5. Valorar la relevancia del sistema informático elaborado por criterio de expertos.

Para el desarrollo de la investigación se utilizaron como **métodos teóricos**:

Análisis y síntesis: se utilizó para procesar la información teórica y empírica sobre la caracterización del problema, la gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, así



como la elaboración de los fundamentos teóricos, las conclusiones parciales y generales de la investigación.

Histórico – lógico: para hacer un estudio del proceso de gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, el surgimiento y evolución de las herramientas utilizadas para apoyar dichos procesos, así como determinar sus principales características.

Modelación: se empleó durante casi toda la etapa de elaboración del sistema, desde su análisis, hasta su implantación, de manera que el conjunto de modelos de la metodología adoptada, describe desde todas las perspectivas posibles el proceso de elaboración de dicho sistema.

Sistémico – estructural: se utilizó para la descomposición del sistema en los diferentes módulos que los componen, en la determinación de sus nexos; así como para la concepción y ensamblaje de los dos componentes principales del sistema: el componente transaccional, y el componente analítico.

Los **métodos empíricos** que se utilizaron fueron:

Método de observación científica: para la corroboración empírica de las deficiencias detectadas relacionadas con la gestión de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba; así como para la implantación y valoración del sistema elaborado.

Método de entrevista: (no estructurada) para obtener información, así como elementos de análisis del sistema fundamentalmente por parte de expertos en el dominio de aplicación.

Criterio de expertos: se utilizó para valorar y buscar consenso de los expertos seleccionados respecto al sistema elaborado, y constatar su relevancia a través de cada uno de los aspectos puestos a consideración de los mismos. Específicamente se utilizó el Método Delphi.

Método de encuesta: este método se utilizó fundamentalmente para la valoración del sistema por criterios de expertos.



Revisión de documentos: para estudiar soluciones antes propuestas y las vías por la que se puede favorecer la gestión de la información relacionada con el sistema de referencia. Se utilizó la revisión de artículos y libros especializados.

Por otra parte se utilizó la **Estadística Descriptiva** con el objetivo de procesar los resultados obtenidos en la consulta a los expertos.

La novedad científica del presente trabajo radica en la obtención de un software que soporta la aplicación de técnicas para el procesamiento transaccional y analítico en línea, enfocado a la explotación de grandes volúmenes de datos del Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba.

Este software permite el análisis, interpretación y presentación de la información que soportan la toma de decisiones en el orden epidemiológico en los distintos niveles del sector de Salud en la provincia. El sistema informático obtenido proporciona un conjunto de herramientas que permiten la evaluación y predicción de tendencias de esta enfermedad desde diferentes perspectivas de análisis así como su trasfondo epidemiológico en el territorio. La utilización de herramientas matemático – computacionales hace mucho más dinámico el proceso de toma de decisiones en el dominio de aplicación.

El documento está conformado por Introducción, Capítulo 1, Capítulo 2, Conclusiones, Recomendaciones, Bibliografía y Anexos.

El Capítulo 1 ofrece en síntesis cada uno de los aspectos fundamentales para construir el marco teórico en el que se sustenta la investigación. En él se abordan temas relacionados con el objeto de estudio y el campo de acción de la investigación, así como un estudio bibliográfico actualizado sobre las principales tendencias y tecnologías actuales, incluyendo las herramientas usadas en el desarrollo de aplicaciones para el procesamiento transaccional y analítico en línea.

El Capítulo 2 abarca todo el proceso de ingeniería y gestión del software, incluyendo la modelación del negocio, la captura de requerimientos, la modelación, diseño e implementación del sistema, además de la valoración del sistema elaborado por criterio de expertos.



CAPÍTULO 1. FUNDAMENTOS TEÓRICOS DEL SISTEMA DE VIGILANCIA EPIDEMIOLÓGICA DEL DENGUE Y HERRAMIENTAS PARA LA AUTOMATIZACIÓN DE SU GESTIÓN

En el presente capítulo se hace un análisis de teorías, conceptos y definiciones, que sirven de base para comprender el objeto de estudio, así como su contexto actual. De ahí que se aborden temas relacionados con la fundamentación del tema tratado, el análisis de otras soluciones, las principales tendencias y tecnologías actuales que sustentan el desarrollo de esta investigación. Además, se ofrece una valoración de las herramientas utilizadas para la confección del software propuesto.

1.1. El Sistema de Vigilancia Epidemiológica del Dengue

El concepto de vigilancia en salud ha sido incorporado al sistema de salud cubano y materializado en cada uno de los niveles de atención [MINSAP, 1995].

Los sistemas de vigilancia epidemiológica exigen un conjunto de condiciones para su desarrollo e implantación. En primer lugar es necesaria la formación y disponibilidad de recursos humanos calificados para el desarrollo de las actividades de vigilancia. Además es necesaria la creación de un sistema de notificación estable de los eventos de salud, así como el hecho de poder contar con herramientas que ayuden a ejecutar tareas de una forma eficaz, formular recomendaciones y tomar decisiones adecuadas.

Un sistema de vigilancia epidemiológica debe ser capaz de:

1. Detectar oportunamente patrones no usuales de los eventos de salud.
2. Cuantificar el impacto de la ocurrencia de esos eventos.
3. Evaluar los mecanismos que se usan en el propio sistema de vigilancia.

Para el alcance de estos objetivos se propone la aplicación combinada de:



- Métodos estadísticos y herramientas que permitan analizar los eventos que se observan de manera sistemática y rutinaria por los especialistas a cargo de la vigilancia.
- Investigaciones epidemiológicas que posibiliten conocer las causas que generan los problemas de salud, así como la relación salud – enfermedad en los territorios y grupos de población.
- Vinculación con las fuentes de información o la comunidad que se vigila.

A los efectos de considerar los datos necesarios para la vigilancia, se toma como referencia un conjunto de fuentes de donde pueden ser obtenidos. Estas pueden describirse como: registros, censos y encuestas demográficas, registros de estadísticas vitales, registros de morbilidad de hospitales y de centros de salud ambulatorios, registros de unidades de salud ocupacional e industrias, registro de seguridad social, registros y reportes de laboratorios de diferentes tipos [Rodríguez, 1999].

En la provincia de Santiago de Cuba, es el Laboratorio SUMA¹ el encargado de realizar los procedimientos para la ejecución de la prueba de detección de anticuerpos IgM a los cuatro serotipos del virus del Dengue en suero mediante la tecnología "SUMA", en todo el territorio; ensayo clínico que permite la confirmación de la enfermedad del dengue.

Las muestras son tomadas en cada uno de los centros asistenciales de la provincia, conservadas y enviadas a este laboratorio. Una vez recibidas se les realiza este proceder denominado UMELISA Dengue IgM Plus.

La validación, interpretación de resultados y su impresión son efectuadas automáticamente por el equipo SUMA, con el programa UMELISA DENGUE IgM Plus, creado a tal efecto.

Toda muestra se procesará nuevamente dos veces por duplicado para definir la conducta a seguir. En caso de ser positiva, se procede a informar a la persona

¹ Dependencia del Departamento de Microbiología del CPHE de Santiago de Cuba.



como “en estudio” o “confirmada”, y se envía al laboratorio de referencia del dengue en el Instituto de Medicina Tropical Pedro Kouri (IPK) en La Habana.

Una vez realizados los ensayos, se procede al registro de cada determinación. De estas se registrarán por las diferentes áreas de salud del territorio, de qué muestra fue hecha la(s) determinación(es), fecha en que fue realizada, resultado de la(s) misma(s), así como a qué paciente pertenece la muestra utilizada.

Estos datos son informados al Departamento de Estadística, el cual consolida la información en modelos, donde presentan la información de forma resumida por meses, trimestres y años. Los indicadores que se incluyen en los mismos son: cantidad de muestras, cantidad de determinaciones, determinaciones positivas, 10% de las determinaciones negativas enviadas al IPK, kits utilizados para el diagnóstico.

Este consolidado es luego enviado a los directivos del CPHE y a la Dirección Provincial de Salud Pública en Santiago de Cuba, para su análisis y la consecuente toma de decisiones.

1.2. Automatización del Sistema de Vigilancia Epidemiológica del Dengue

El desarrollo tecnológico en general y el de la informatización en particular inciden positivamente de manera acelerada en la Salud, sector de elevada complejidad con necesidades sentidas y reales relacionadas con la adquisición de nuevos conocimientos de manera permanente. Estas necesidades se ponen a relieve ante los constantes retos de la salud humana, así como por las transformaciones organizacionales que ocurren en los sistemas y servicios de salud.

En virtud de elevar la calidad de los servicios de salud, se han diseñado e implementado variantes para automatizar este proceso. Pasos importantes han sido dados desde el mismo inicio de la Revolución Cubana. Se expresa, como eje central de esta voluntad, la *Estrategia de Informatización del Sector de la Salud Pública*, de acuerdo con la *Política de Informatización de la Sociedad Cubana*. Esta define como parte de los elementos importantes a tener en cuenta, el hecho de fomentar la expansión de proyectos que agilicen y hagan más eficientes los



procesos de trámites y fácil acceso a la información; así como impulsar el desarrollo de aplicaciones de servicios de información fomentando el uso de las tecnologías Web.

La *Estrategia para la Informatización de la Salud Pública* ha recibido la correspondiente influencia de esta política de informatización, significando las áreas de: Atención Primaria (Policlínicos y Clínicas Estomatológicas), Hospitales, Vigilancia Epidemiológica (Vigilancia en Salud), Sistema Económico, Formación y Perfeccionamiento de Recursos Humanos, entre otros.

La introducción del equipamiento y los programas de computación a nivel internacional ha permitido a los profesionales de la Salud Pública obtener, en condiciones ordinarias, resultados en vigilancia mucho más eficientes. Además ha dado autoridad a sus profesionales para organizar, comunicar, tabular y analizar datos. En este sentido, su uso ha incrementado la oportunidad en la recolección y análisis de datos, y ha disminuido la dependencia de los epidemiólogos sobre los programadores y bioestadísticos para el análisis e interpretación de datos.

Dada la voluntad política y la necesidad de asegurar la calidad y operatividad en el procesamiento y análisis de la información, se estableció una Red Nacional de Vigilancia en Salud. Esta tiene como propósito garantizar el intercambio estable de la información entre las diversas unidades del Sistema Nacional de Salud que intervienen en la vigilancia.

Esta red ha operado a partir de la instalación de una red de computadoras en la Unidad Central de Análisis y Tendencias en Salud, que cuenta con estaciones de trabajo en las principales Direcciones Nacionales del Área de Higiene y Epidemiología, y la Dirección de Estadísticas del MINSAP. A la misma se conectan todas las unidades provinciales del país, algunos de los institutos de investigaciones relacionados con esta área, así como el resto de las unidades de salud integradas a la red.

Lograr una adecuada y homogénea explotación de los recursos de computación en todo el país constituyó desde el principio una premisa fundamental en la



implementación del sistema. Para ello se distribuyó una metodología de trabajo, un gran volumen de programas y paquetes computacionales de actualidad y amplio uso en el trabajo epidemiológico a nivel internacional, los utilitarios de uso general y la bibliografía correspondiente. Paralelamente se impartieron cursos a los especialistas responsables de su explotación.

En la actualidad se ha fortalecido de forma decisiva la estructura a nivel provincial, con la instalación de redes locales en los centros provinciales de Higiene y Epidemiología. Estas permiten vincular armónicamente las subdirecciones que las integran. Por otra parte brindan la posibilidad de ampliarse hacia algunos municipios del país, lo que incrementa considerablemente la cobertura de la red nacional, que constituye el soporte para el procesamiento, análisis y transmisión de la información para la acción.

El desarrollo alcanzado en esta esfera en este período ha permitido, más que automatizar simplemente una tarea o un grupo de operaciones manuales, lograr cambios en los métodos de análisis por los profesionales epidemiólogos en general. Esto obviamente da respuesta a las necesidades propias que el sistema de salud reclama a cada nivel, lo que ha constituido una incuestionable ventaja.

Se tiene referencia de la existencia del Sistema Automatizado de Vigilancia en Salud (EpiAlerta). Este tiene como objetivo fundamental la recolección y transmisión de la información, que permite viabilizar todas las acciones que se realizan en el sistema alerta – acción a cada uno de los niveles. Sin embargo es muy general, es decir, ha sido concebido para todas las enfermedades infecto – contagiosas y genera salidas preestablecidas a través de consultas tradicionales.

Por ello se propone en esta investigación la creación de un sistema informático que permita el procesamiento transaccional y analítico en línea, de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, que tribute al proceso de toma de decisiones administrativas en el orden epidemiológico, a través de técnicas y herramientas que flexibilicen el análisis de los directivos.



1.3. Sistemas de apoyo a la toma de decisiones

Los sistemas de información están cambiando la forma en que operan las organizaciones actuales. A través de su uso se logran importantes mejoras, pues automatizan los procesos operativos de las empresas, proporcionan información de apoyo al proceso de toma de decisiones y, lo que es más importante, facilitan el logro de ventajas competitivas a través de su implantación en las empresas [Cohen, 1996].

Estos pueden ser sistemas transaccionales o sistemas de apoyo a la toma de decisiones. Los sistemas de información que logran la automatización de procesos operativos dentro de una empresa son llamados sistemas transaccionales. Los sistemas que apoyan el proceso de toma de decisiones son llamados sistemas de soporte a la toma de decisiones.

En un sentido más amplio, los sistemas de apoyo a la toma de decisiones son un conjunto de programas y herramientas que permiten obtener de manera oportuna la información que se requiere durante el proceso de la toma de decisiones que se desarrolla en un ambiente de incertidumbre. Además proporcionan la mayor cantidad de información relevante en el menor tiempo posible, con el fin de decidir lo más adecuado y oportuno. De ahí que la información adquiera un papel imprescindible para el apoyo a dicho proceso [McClure, 1978].

Los sistemas de apoyo a la toma de decisiones pueden agruparse en cuatro grandes grupos, en dependencia de su finalidad y recursos utilizados para lograr la implementación de sus funcionalidades:

- Sistemas de Soporte a la Toma de Decisiones (DSS del inglés *Decision Support Systems*) que tienen como finalidad apoyar la toma de decisiones mediante la generación y evaluación sistemática de diferentes alternativas o escenarios de decisión, todo esto utilizando modelos y herramientas computacionales.
- Sistemas de Información para Ejecutivos (EIS del inglés *Executive Information Systems*), los cuales están dirigidos a apoyar el proceso de



toma de decisiones de los altos ejecutivos de una organización, presentando información relevante y usando recursos visuales y de fácil interpretación, con el objetivo de mantenerlos informados.

- Sistemas para la Toma de Decisiones de Grupo (GDSS del inglés *Group Decision Support Systems*), los cuales cubren el objetivo de lograr la participación de un grupo de personas durante la toma de decisiones en ambientes de anonimato y consenso; apoyando decisiones simultáneas.
- Sistemas Expertos de Soporte a la Toma de Decisiones (EDSS del inglés *Expert Decision Support System*), los cuales permiten cargar bases de conocimiento que se integran por una serie de reglas de sentido común para que diferentes usuarios las consulten, apoyen la toma de decisiones, la capacitación, etc. [Cohen, 1996].

Como sistemas informáticos, consisten habitualmente en varios componentes: bases de datos fuentes, sistemas de extracción – transformación – carga de datos (ETL del inglés *Extraction – Transformation – Load*), almacenes de datos (DW del inglés *Datawarehouse*), herramientas de procesamiento analítico en línea (OLAP del inglés *Online Analytic Processing*), bases de datos multidimensionales y otras herramientas de análisis de información [Vega, 2008]. La mayor parte de estos componentes serán analizados con más detalles en los epígrafes siguientes.

1.4. Inteligencia de negocio para el apoyo a la toma de decisiones

La adopción de modelos computacionales orientados a apoyar el proceso de toma de decisiones ha cobrado una fuerza extraordinaria en las últimas décadas. Esto se debe a que dichos modelos tienen la finalidad de representar, mejorar y ofrecer nuevas alternativas en cuanto al desempeño y actuación de las organizaciones. Además alcanzan mayor relevancia cuando se cuenta con un volumen de información considerable, y que puede ser utilizado como la fuente primaria para generar escenarios, pronósticos, evaluaciones y balances que apoyen la toma de decisiones. Este proceso es conocido como Inteligencia de Negocio (BI del inglés *Business Intelligence*).



Nader define la inteligencia de negocio como aquellas tecnologías que en su conjunto permiten a las organizaciones utilizar información disponible de cualquier parte de ella, en función de realizar análisis adecuados, descubrir nuevas oportunidades, y en consecuencia tomar mejores decisiones [Nader, 2006].

Moss por su parte plantea que la inteligencia de negocio es la combinación de técnicas y herramientas, que posibilitan aprovechar la información estratégica subyacente en los datos operacionales generados a diario y acumulados de cualquier organización. Esta combinación tiene como propósito fundamental el apoyo al proceso de toma de decisiones [Moss, 2003].

La inteligencia de negocio permite analizar datos acumulados en una organización y extraer una cierta inteligencia o conocimiento de ellos. Varios son los elementos que componen este proceso, entre los que se pueden mencionar la multidimensionalidad, que indica la capacidad de resumir información dispersa, relacionada o no desde diferentes perspectivas de análisis; y la instrumentación de almacenes de datos, que indica la capacidad de buscar y analizar la información.

La utilización de la inteligencia de negocio permite identificar patrones válidos, novedosos y potencialmente útiles, que a través de técnicas tradicionales como por ejemplo el lenguaje de consultas estructurado (SQL del inglés *Structured Query Language*), sería imposible o difícil de obtener. Normalmente por técnicas tradicionales se obtiene aproximadamente el 80% de la información; que la mayoría de las veces no contiene la información más relevante y estratégica para una organización.

Desde un punto de vista más pragmático, y asociándolo directamente a las tecnologías de la información, se puede definir la inteligencia de negocio como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información no estructurada, en información estructurada, para su explotación directa o su análisis y conversión en conocimiento de apoyo a la toma de decisiones sobre el negocio [Vega, 2008].

Estas herramientas y metodologías tienen en común las siguientes características:



- *Accesibilidad a la información:* los datos constituyen la fuente principal de este concepto. Lo primero que debe garantizar este tipo de herramientas y técnicas será el acceso a los datos con independencia de su procedencia.
- *Apoyo a la toma de decisiones:* más que presentar información, se busca que los usuarios finales tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen y les sean útiles para tomar cualquier decisión.
- *Orientación al usuario final:* se busca independencia entre los conocimientos técnicos de los usuarios y su capacidad para utilizar estas herramientas.

En la Figura 1.1 se muestran las cinco fases en que Ricardo D. Bernabeu divide el proceso de inteligencia de negocio [Bernabeu, 2009].

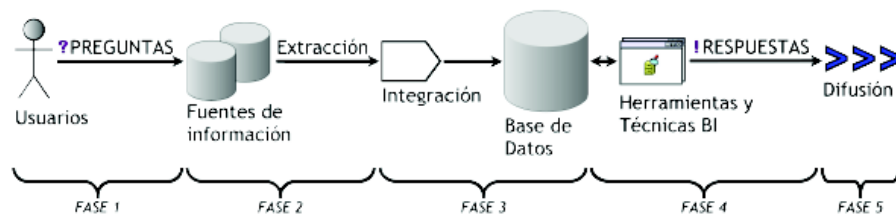


Figura 1.1. Proceso de Inteligencia de Negocio. Fuente: [Bernabeu, 2009]

Fase 1 – Dirección y planeación. Es en esta fase donde se deberá precisar los requerimientos y necesidades de información cada usuario, para luego con su apoyo generar las preguntas que ayudarán a alcanzar sus objetivos.

Fase 2 – Recolección de información. Una vez definidos los requerimientos, necesidades, y preguntas, se obtienen los datos necesarios para responder a las mismas, y cubrir dichas necesidades y requerimientos. Estos datos son extraídos desde las diferentes fuentes de información, que pueden ser internas o externas a la organización.

Fase 3 – Procesamiento de datos. En esta fase se lleva a cabo un proceso de integración y carga de los datos en un formato útil para el análisis. Esta actividad



puede realizarse mediante la creación de una nueva base de datos, agregando datos a una base de datos ya existente o bien consolidando la información.

Fase 4 – Análisis y producción. Ya en esta fase se trabaja directamente sobre los datos extraídos e integrados, utilizando herramientas y técnicas propias de la tecnología de inteligencia de negocio. Con la creación de vistas de análisis, reportes, cuadros de mando, gráficos estadísticos, entre otros, se obtendrán las respuestas a las preguntas formuladas en la fase inicial.

Fase 5 – Difusión. En esta última fase, los usuarios podrán explorar los datos de una manera amena, sencilla e intuitiva, a través de las herramientas pertinentes a tal efecto.

1.4.1. Almacén de datos o datawarehouses

Ante la necesidad de resumir grandes volúmenes de datos para obtener información útil y relevante para el análisis, se implementan estructuras que más que eficiencia en la representación de los datos, brinden mejores prestaciones para la presentación y análisis de la información. Dentro de estas estructuras están los almacenes de datos.

Un almacén de datos es una base de datos de sólo lectura. Esta se crea como soporte al proceso de toma de decisiones, que contendrá información histórica de la organización [Imhoff, 2003]. Los datos contenidos en él, se preparan para ser analizados en línea, y no para procesar transacciones en tiempo real como suele ocurrir en los sistemas de proceso de transacciones en línea (OLTP del inglés *Online Transactional Processing*).

Esta tecnología relativamente reciente, tiene como propósito ofrecer metodologías para recopilar e integrar los datos históricos de una organización; además de organizar grandes volúmenes de datos de procedencia generalmente estructurada [Imhoff, 2003].

Este tipo de persistencia es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma, ambiente muy diferente a los sistemas operacionales.



Bill Inmon y Ralph Kimball, considerados los padres de los almacenes de datos, definieron sus propios paradigmas en este campo, y que toman sus nombres:

El paradigma de Inmon

Una organización debe tener un DW y varios mercados de datos² que se nutran de la información del primero. En este, la información puede estar almacenada en tercera forma normal. Bill Inmon definió como características fundamentales de un DW las siguientes:

- *Integrado*: los datos almacenados tienen que estar integrados en una estructura consistente.
- *Temático*: sólo los datos que sean necesarios para el proceso de generación del conocimiento se integrarán desde el entorno operacional.
- *Histórico*: el tiempo es una parte implícita de la información contenida en un DW. En los sistemas operacionales, los datos van a mostrar siempre el estado de la actividad del negocio en el momento en que se consulten. Contrario a esta realidad, la información almacenada en un DW sirve para realizar análisis de tendencias, comparaciones, etc. Esto sólo es posible cargando los distintos valores que toman las variables de análisis en diferentes instantes de tiempo.
- *No volátil*: un DW puede ser leído, pero no modificado. La información es por tanto permanente. Esto implica que la única acepción de actualizar un DW se refiera a la incorporación de los últimos valores que tomaron las distintas variables contenidas en él, sin ningún tipo de acción sobre lo que ya existía [Inmon, 2002].

El paradigma de Kimball:

Contrario al paradigma de Inmon, Kimball plantea que un DW es el resultado de la unión de los mercados de datos de las diferentes áreas de una organización; es

² Versión especial de almacén de datos. Son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones.



decir, que su construcción se lleva a cabo por la integración de varios mercados de datos [Kimball, 2002].

Ambos paradigmas son válidos pero se considera al Kimball como el más implementado en la actualidad, dado que la mayoría de las organizaciones por diversos motivos (entre los que se destacan el tiempo y el costo de producción), comienzan por la implementación de varios mercados de datos que posteriormente se integran en un DW.

1.4.2. Elementos de un modelo multidimensional

Como se hizo mención en el Epígrafe 1.4, la multidimensionalidad es uno de los componentes fundamentales en la inteligencia de negocio, específicamente en los almacenes de datos. De ahí que en el diseño de este tipo de estructuras, la el diseño de un modelo multidimensional (MMD) adquiera gran relevancia, encontrando su fundamento en el modelo entidad – relación (MER), y en la ingeniería de texto y datos numéricos

Un MDD tiene como ventajas sobre el MER, que es flexible, está desnormalizado y orientado a los intereses de un usuario final. Lo anterior no implica que existan inconsistencias en los datos contenidos en el modelo. Además se disminuye el número de tablas y relaciones entre ellas, lo que permite agilizar el acceso a los datos [Vega, 2008].

El MDD se implementa a través de dos estructuras fundamentales: las tablas de dimensiones y las tablas de hechos.

- *Tablas de hechos:* a través de estas estructuras, se representa la ocurrencia de un determinado proceso dentro de una organización, que pueden o no guardar relación entre sí. Por lo general, almacenan medidas numéricas que representan valores de las dimensiones a las que están asociadas. La llave de este tipo de estructura es una llave compuesta, formada por la composición de las llaves primarias de las tablas de dimensionales a las que está relacionada.



- *Tablas de dimensiones:* a través de estas estructuras, se representa la ocurrencia de las diferentes perspectivas de análisis, y contienen una llave simple y uno o varios atributos que la describen, y eventualmente llaves foráneas de otras tablas de dimensión, en dependencia del esquema de diseño que se asuma. Los atributos dimensionales son fundamentalmente textos descriptivos, y juegan un papel determinante al servir como restricciones en la mayoría de las consultas que realizan los usuarios.

Con relación a esta última estructura es válido destacar que existe una dimensión fundamental en todo DW: la dimensión tiempo. Esto se debe a que todo registro que se incluya, constituye la ocurrencia de un fenómeno en un instante de tiempo definido. Por ello se afirma que la dimensión de referencia es la que establece uno de los objetivos fundamentales de la construcción de un DW: la conservación de un “histórico”.

La forma en que las estructuras anteriormente descritas se combinen y se relacionen, determinará el esquema que adoptará el MMD [Date, 2003]. Existen varios esquemas para el modelado de los datos en un DW, siendo los más utilizados:

- *Esquema en estrella:* en este esquema la tabla de hechos está en el centro de la estrella, y de forma radial se relacionan con ella todas las tablas de dimensiones, las cuales no se relacionan entre sí.
- *Esquema de copo de nieve:* constituye una extensión del esquema anteriormente descrito, con la única diferencia que existen jerarquías en las dimensiones. La ventaja de este esquema, es que ocupa menor espacio de almacenamiento; sin embargo, incrementa proporcionalmente la complejidad de las consultas con el aumento del número de tablas con las que se debe interactuar.
- *Esquema de constelación:* en este esquema existe una tabla principal de hechos, una o más tablas auxiliares de hechos, y varias tablas de dimensiones; por lo que se dice que este esquema se forma por la unión de



una serie de esquemas en estrella. Este esquema no es soportado por todas las herramientas de consulta y análisis.

1.4.3. Procesamiento Analítico en Línea

El Procesamiento Analítico en Línea, concepto que fue mencionado en el Epígrafe 1.3, posibilita a los usuarios utilizar de una forma cómoda, intuitiva, y natural los DW para el análisis en línea de los datos, ofreciendo a los mismos un estilo de “navegación” sobre estos. En tal sentido se pueden obtener respuestas rápidas a consultas analíticas complejas e interactivas, acelerando la entrega de información a los usuarios finales que ven estas estructuras de datos como cubos multidimensionales debido a que la información es vista desde varias dimensiones.

La entrega a la que se hace referencia en el párrafo anterior, es optimizada al realizar algunos cálculos por adelantado, en vez de realizar el cálculo al momento de la solicitud. Esta característica permite analizar información de una forma más rápida y eficiente que lo que es posible con tecnologías de bases de datos relacionales solamente.

Dentro de la utilización de técnicas de almacenes de datos, el componente OLAP es un elemento primordial para su utilización y explotación. Este proporciona la funcionalidad esencial para un conjunto de aplicaciones y herramientas que incluyen desde reportes corporativos, hasta el soporte avanzado a la toma de decisiones [Pence, 2008].

Para los sistemas OLAP, el análisis multidimensional de datos corporativos es una característica necesaria. Esta posibilita la navegación por los datos, reflejada en la visualización de la estructura multidimensional, en unos campos de selección que permitan elegir el nivel de agregación (jerarquía) de la dimensión, y/o la elección de un dato en concreto, etc.

Por otra parte, el procesamiento analítico de los DW puede ser esencialmente de dos tipos: OLAP Multidimensional (MOLAP), y OLAP Relacional (ROLAP).

En la tecnología MOLAP, los datos se almacenan en cubos multidimensionales, lo



que posibilita una rápida recuperación de datos, elevando así el rendimiento.

Por su parte, la tecnología ROLAP supone la manipulación de datos almacenados en bases de datos relacionales, ofreciendo la apariencia de la funcionalidad de pivoteo de un OLAP tradicional. En esencia, todas las acciones de dicha funcionalidad son equivalentes a añadir cláusulas WHERE en las sentencias SQL. Esto como es lógico acarrea un bajo rendimiento en las prestaciones, al ser excesivamente alto el tiempo de consulta y recuperación de datos. Otro de sus inconvenientes se relaciona con las limitaciones de las funcionalidades del SQL, ya que las mismas no proveen todas las estructuras necesarias para este tipo de análisis.

1.5. Tendencias y tecnologías actuales de desarrollo

En esta investigación fue necesario realizar un análisis sobre las tendencias y tecnologías actuales. A continuación se resumen las utilizadas en el presente trabajo.

1.5.1. Modelo Cliente – Servidor

Durante la investigación presentada se descartó la posibilidad de adoptar la alternativa de implementar una aplicación estándar como posible solución a la problemática del CPHE de Santiago de Cuba. Con ese tipo de aplicación no existe la posibilidad de compartir los datos almacenados, ni los resultados de cada análisis o procesamiento que se haga con ellos. Además, debería estar instalada la misma aplicación en cada una de las computadoras del personal encargado del control y análisis de este sistema, lo que implicaría duplicidad de la información, inconsistencia de los datos y la no veracidad de la información.

Es por ello que se tuvo en cuenta el modelo cliente – servidor; teniendo como alternativas de solución la instrumentación de una aplicación cliente – servidor tradicional, o una aplicación cliente – servidor con navegador.

Este modelo, en el que el procesamiento es compartido entre computadoras que funcionan como clientes y otras que funcionan como servidores, se logra concentrar la información, logrando consistencia e integridad de la misma [Harkey,



1996]. Por otro lado se logra mayor rapidez en la ejecución de ciertas operaciones (transacciones, bloqueos, actualizaciones, etc.), y se garantiza que cualquier inversión en equipamiento de los servidores sea de beneficio inmediato para el resto de las computadoras clientes, con la consiguiente disminución de costos.

Para utilizar una aplicación cliente – servidor tradicional se necesita desarrollar una aplicación cliente para visualizar la información. Este tipo de aplicación tiene dos desventajas notables: todas las computadoras en las que los usuarios necesiten el sistema deberán tener instalada esta aplicación. Además, para desarrollar el mantenimiento a la aplicación cliente se requiere actualizar en todas las computadoras en que esté instalada. Esto puede requerir mucho tiempo de desarrollo y también para la actualización.

Por otra parte, una aplicación cliente – servidor no tradicional utiliza un navegador para visualizar la información. Esto posibilita una reducción del tiempo de desarrollo pues sólo es necesario elaborar la aplicación que se instalará en el servidor. También desaparece el tiempo para la instalación del sistema en las computadoras clientes. La mayoría de los sistemas operativos tienen incluido un navegador que se puede utilizar para visualizar la información. El mantenimiento a una aplicación de este tipo solo se realiza en el servidor, donde se encuentra toda la información, por lo que no es necesario modificar la aplicación cliente ni realizar instalaciones adicionales [Harkey, 1996].

Dadas las características de los modelos de aplicaciones analizadas y del proceso de registro, control, análisis, interpretación y presentación de la información relacionada con este sistema en el CPHE de Santiago de Cuba, se decidió la instrumentación de una aplicación cliente – servidor con navegador para desarrollar el sistema.

Se analizó el equipamiento básico indispensable para instalar una aplicación cliente – servidor con un navegador es un servidor Web, conexión a red de datos y computadoras clientes. Para el correcto funcionamiento de una aplicación bajo este modelo, con un navegador se requiere de los siguientes elementos de



software: protocolo de comunicación, navegador, gestor de base de datos, y software del servidor Web [Harkey, 1996].

1.5.2. Tecnologías de inteligencia de negocio de código abierto

El uso de las técnicas y arquitecturas de inteligencia de negocio permite solucionar uno de los problemas más comunes que afrontan las organizaciones actuales: el procesamiento de grandes volúmenes de datos y la extracción de conocimiento. Su adopción permitiría sentar las bases para la toma de decisiones estratégicas que influyan directamente en el mejoramiento del desempeño de cualquier organización.

La comunidad “*Open Source*” hoy en día ha incursionado prácticamente en todas las áreas de la Informática y existen algunas donde su supremacía es indudable. En el área de la BI también se ha producido un despegue en el desarrollo de soluciones [Vega, 2008].

Una solución basada en inteligencia de negocio, de acuerdo a sus características debe estar soportada por un conjunto de herramientas donde se establece una cooperación entre ellas. Dichas herramientas permitirían transitar por las diferentes etapas del proceso de análisis de los datos, desde la adquisición hasta la visualización de los resultados. Entre las principales herramientas disponibles están:

- Herramientas ETL: Kettle, Clover, Enhydra Octopus, Data Integration, etc.
- Desarrollo OLAP: Mondrian y JPivot.
- Minería de Datos: WEKA, YALE y otras herramientas con versiones libres limitadas como: Tiberius, WizWhy, CART y See5 / C5.0.
- Motores de Reportes y Gráficos: JFreeReport, BIRT, JasperReport y JFreeChart.
- Entorno de desarrollo para Cuadros de Mando (*Dashboards*): JetSpeed y JBoss Portal.
- Gestores de Bases de Datos: MySQL y PostgreSQL.
- Soluciones completas: Pentaho y SpagoBI.



1.5.3. Pentaho Open Source Business Intelligence

La comunidad Open Source ha dado grandes pasos en el desarrollo de herramientas informáticas de inteligencia de negocio. En estos momentos, cuenta con proyectos estables, que constituyen variantes de solución que merecen ser consideradas. Destaca entre ellos, el proyecto *Pentaho Open Source Business Intelligence* que hoy integra el conjunto de herramientas necesarias para dar una respuesta informática coherente, confiable y escalable a las necesidades de información de las organizaciones modernas.

Pentaho ofrece soluciones en áreas tales como: análisis de información, reportes, tableros o cuadros de mando³, flujos de trabajo, minería de datos, procesos ETL, planificador, seguridad, entre otras; pretendiendo ser una alternativa de código abierto, ante soluciones propietarias tradicionales, entre las que se destacan: Business Objects, Cognos, Microstrategy, Microsoft, etc.

Algunas de sus características son:

- Plataforma 100% J2EE, lo que asegura su escalabilidad, integración y portabilidad.
- Servidor: puede correr en servidores compatibles con J2EE como JBOSS, WebSphere, Tomcat, WebLogic y Oracle, lo que posibilita que toda la información pueda ser accedida mediante un navegador.
- Base de datos: vía JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird.
- Sistema operativo: no hay dependencia de uno u otro. Lenguaje interpretado.
- Lenguaje de programación: Java, Javascript, JSP, XSL (XSLT/XPath/XSL-FO).
- Interfaz de desarrollo: Java SWT, Eclipse, basada en tecnología Web.
- Repositorio de datos basado en XML.

³ Conocidos en inglés como *dashboards*.



Las soluciones Pentaho brindan una infraestructura de herramientas de análisis e informes, integrados con un motor del flujo de los procesos de negocio, capaz de ejecutar las reglas de negocio, expresadas en forma de procesos y actividades.

De todas estas soluciones disponibles en la plataforma Pentaho, se decidió utilizar las siguientes:

Mondrian Schema Workbench

Es un entorno visual para el desarrollo y prueba de cubos OLAP Mondrian. Su propósito es la creación y publicación de esquemas que definen cubos multidimensionales utilizando el lenguaje XML. Si bien la definición del XML para esquemas Mondrian no es extremadamente compleja, en la práctica resulta engorrosa su utilización. Con esta aplicación, se puede establecer una conexión JDBC como el modelo físico, para luego elaborar el esquema lógico de manera simple y efectiva [Pentaho, 2008].

Para su utilización primeramente se deben especificar la fuente de los datos y el cubo OLAP en el fichero de configuración XML: *Mondrian Cube Schema*. En este fichero se pueden definir las dimensiones, los niveles de jerarquía de dimensiones, los hechos y conexión a la base de datos.

Pentaho Data Integration

Solución de Pentaho que incluye un conjunto de herramientas para realizar las tareas ETL (extracción, transformación y carga). Su objetivo principal es que el proyecto sea fácil de generar, mantener y desplegar. Este tiene cuatro herramientas:

- SPOON: permite diseñar de forma gráfica la transformación ETL.
- PAN ejecuta las transformaciones diseñadas con SPOON.
- CHEF permite, mediante una interfaz gráfica, diseñar la carga de datos incluyendo un control de estado de los trabajos.
- KITCHEN permite ejecutar los trabajos batch diseñados con CHEF.



Pentaho Design Studio

Herramienta de desarrollo (IDE del inglés *Integrated Development Environment*) que constituye una versión de Eclipse integrada a Pentaho. Proporciona una amplia gama de funcionalidad al usuario como: administración de proyectos, las plantillas del proyecto, y una interfaz del usuario totalmente personalizable, entre otras. Además incorpora un conjunto de características con las que se pueden crear aplicaciones Java, incluidas aplicaciones J2EE, de capa Web e independientes.

Además, Pentaho Design Studio incluye un módulo que permite la creación de elementos de inteligencia de negocio: *xactions*, consultas, esquemas XML, etc.; lo que deja ver la factibilidad de esta herramienta para el desarrollo de este tipo de aplicaciones.

Pentaho Report Designer

Herramienta de la plataforma que ofrece la solución adecuada a las necesidades de los usuarios para la generación de reportes. Esta es una solución basada en el proyecto JFreeReport, que permite crear, publicar y generar de forma ágil diferentes informes. Permite además la distribución de los resultados del análisis en múltiples formatos, incluyendo la opción de imprimir o exportar a formato PDF, XLS, HTML y texto. Los reportes Pentaho permiten también la programación de tareas y ejecución automática de informes con una determinada periodicidad.

Apache Tomcat

En la actualidad se dispone de una multitud de servidores Web de código libre, entre ellos: JavaServer Web Development Kit, Enhydra, Apache Tomcat, etc.

Muchas son las ventajas que ofrecen unos sobre otros, sin embargo, Tomcat es la implementación de referencia oficial para las especificaciones Servlet y JSP (elementos implementados y utilizados en la plataforma Pentaho), a partir de las versiones 2.2 y 1.1 respectivamente [Mateu, 2004]. Además es el servidor que viene incluido en la plataforma de Pentaho.



Tomcat provee una implementación por defecto de la interfaz del servidor y raramente es personalizada por los usuarios. Es posible instalar este servidor como un servicio para comodidad de quienes lo utilizan. Su único punto débil reside en la complejidad de su configuración, dado el gran número de opciones existente.

1.5.4. Proceso Unificado de Desarrollo de Software

Luego de realizar una revisión bibliográfica de las metodologías de desarrollo de software (ICONIX, SCRUM, y varias metodologías ágiles), analizar sus características, y cuanto se ajustaban a las necesidades de desarrollo y documentación del sistema a elaborar, se decidió adoptar la metodología del Proceso Unificado de Desarrollo de Software.

Este es un proceso de desarrollo de software para la ingeniería de software orientado a objetos. Tiene dentro de sus principales características el hecho de ser iterativa e incremental, que proporciona un enfoque disciplinado para asignar tareas y responsabilidades dentro de una organización de desarrollo de software. Su meta es asegurar la producción de un producto (software) con la más alta calidad, que reúna las necesidades de los usuarios dentro del cronograma planeado y la inversión prevista. Se divide en cuatro fases:

- *Inicio*: el objetivo en esta etapa es determinar la visión del proyecto.
- *Elaboración*: en esta etapa el objetivo es determinar la arquitectura óptima.
- *Construcción*: en esta etapa el objetivo es llevar a obtener la capacidad operacional inicial.
- *Transmisión*: el objetivo es llegar a obtener la distribución del proyecto.

Cada una de estas etapas es desarrollada mediante iteraciones, que consisten en reproducir el ciclo de vida en cascada a menor escala. Los objetivos de una iteración se establecen en función de la evaluación de las iteraciones precedentes [Jacobson, 2000]. Vale mencionar que el ciclo de vida que se desarrolla por cada iteración, es llevada bajo dos disciplinas:



Disciplina de Desarrollo

- *Ingeniería de negocios*: con esta actividad se trata de entender las necesidades del negocio.
- *Requerimientos*: se trasladan las necesidades del negocio a un sistema automatizado.
- *Análisis y diseño*: se trasladan los requerimientos dentro de la arquitectura de software.
- *Implementación*: se crea el software que se ajuste a la arquitectura y que tenga el comportamiento deseado.
- *Pruebas*: actividad en la que se busca comprobar que el comportamiento requerido es el correcto y que todo lo solicitado haya sido implementado.

Disciplina de Soporte

- *Configuración y administración del cambio*: con esta actividad se guardan todas las versiones del proyecto.
- *Administración del proyecto*: actividad en la que se administran horarios y recursos.
- *Ambiente*: se administra el ambiente de desarrollo.
- *Distribución*: tareas que encaminadas a darle salida al proyecto.

Es recomendable que a cada una de estas iteraciones se les clasifique y ordene según su prioridad, y que cada una se convierta luego en una entrega al cliente. Esto trae como beneficio la retroalimentación que se tendría en cada entrega o en cada iteración.

Por último, cabe destacar como una ventaja de esta metodología, el hecho de que hace uso de del lenguaje de modelado UML (del inglés *Unified Modelling Language*). Este lenguaje hace más estándar y entendible a todos los desarrolladores y clientes el proceso de desarrollo [Larman, 2000], debido a que es un lenguaje gráfico para visualizar, especificar, construir y documentar los artefactos de un software; y definir los mecanismos de extensión con la finalidad de cumplir con las necesidades específicas de desarrollo. Además, este es



soportado por la herramienta CASE Rational Rose, que se utilizará para la modelación de todo el proceso de desarrollo.

1.6. Conclusiones del capítulo

En este capítulo se ofrece una visión de los principales conceptos del objeto de estudio. Se abordaron temas relacionados con la metodología, tecnologías y herramientas empleadas para acometer la elaboración del producto informático propuesto.

Para la construcción de este sistema de apoyo a la toma de decisiones en el CPHE de Santiago de Cuba relacionado con el Sistema de Vigilancia Epidemiológica del Dengue, se seleccionó la implementación del mismo como una aplicación cliente – servidor con navegador, con varias soluciones de la plataforma Pentaho Open Source Business Intelligence. Además se escogió el sistema gestor de bases de datos MySQL. No hubo necesidad de incurrir en costo alguno por pago de licencias para la utilización de estas tecnologías, que se distribuyen bajo sistemas de código abierto.



CAPÍTULO 2. CONSTRUCCIÓN DEL SISTEMA DE PROCESAMIENTO TRANSACCIONAL Y ANALÍTICO PARA LA VIGILANCIA EPIDEMIOLÓGICA DEL DENGUE

En este capítulo se muestran las principales características de la solución propuesta, así como su diseño general y valoración; teniendo en cuenta las necesidades que debe cumplir. Para ello se utiliza la metodología RUP y los artefactos necesarios para ilustrar su proceso de desarrollo.

2.1. Breve caracterización del negocio desde sus procesos

Se pudo identificar como el proceso principal del CPHE, el control de la situación higiénico – epidemiológica del territorio. Como proceso de apoyo, que forma parte de este control, está la realización de pruebas de laboratorio con interés higiénico – epidemiológico.

A partir de la identificación de estos dos procesos, cuyas características fueron descritas en el Epígrafe 1.1, se identificaron las reglas del negocio, las cuales rigen el funcionamiento de estos procesos, establecen políticas y restricciones que deben ser cumplidas para llevar a cabo satisfactoriamente los mismos.

Para obtener las reglas del negocio fueron entrevistados varios especialistas en los procesos del negocio que fueron identificados. Entre ellos, especialistas del Laboratorio SUMA, especialistas en Higiene y Epidemiología, especialistas del Departamento de Estadística, así como directivos, todos del CPHE de Santiago de Cuba. Las principales reglas del negocio que se relacionan con la investigación se muestran en el Anexo 1.

Con los procesos del negocio interactúan un conjunto de personas, instituciones y sistemas que se benefician o brindan informaciones a dichos procesos denominados actores. En la Tabla I se muestran los actores del negocio y su justificación.



Tabla I. Actores del negocio

Actor del Negocio	Justificación
Población	Representa a toda persona del territorio, vulnerable a las condiciones higiénico – epidemiológicas del mismo. Se ve beneficiado del control de la situación higiénico – epidemiológica del territorio, así como de todos los procesos de apoyo.

El diagrama de casos de uso del negocio muestra gráficamente la relación existente entre los actores y procesos del negocio. La Figura 2.1 muestra el diagrama de casos de uso del negocio obtenido.

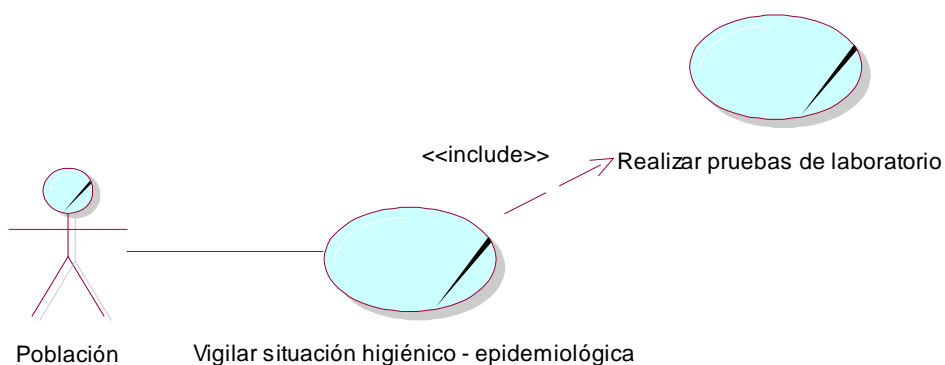


Figura 2.1. Diagrama de casos de uso del negocio

2.2. Requerimientos del sistema

2.2.1. Requerimientos funcionales y no funcionales

Los requerimientos o requisitos funcionales identificados definen el comportamiento que debe tener el sistema a elaborar; corresponden a las acciones, capacidades que el sistema debe realizar o tener [Jacobson, 2000]. Estas actividades a informatizar se muestran en el diagrama de actividades del proceso principal, que aparece en el Anexo 2.

Los requerimientos funcionales que se identificaron fueron los siguientes:



Relacionados con el procesamiento transaccional

- R1. Insertar datos de un paciente.
- R2. Modificar datos de un paciente.
- R3. Eliminar datos de un paciente.
- R4. Visualizar datos de un paciente.
- R5. Buscar datos de uno o varios pacientes según un criterio.
- R6. Ordenar datos de los pacientes según un criterio.
- R7. Insertar datos de una muestra.
- R8. Modificar datos de una muestra.
- R9. Eliminar datos de una muestra.
- R10. Visualizar datos de una muestra.
- R11. Buscar datos de una o varias muestras según un criterio.
- R12. Ordenar datos de las muestras según un criterio.
- R13. Insertar datos de una determinación.
- R14. Modificar datos de una determinación.
- R15. Eliminar datos de una determinación.
- R16. Visualizar datos de una determinación.
- R17. Buscar datos de una o varias determinaciones según un criterio.
- R18. Ordenar datos de las determinaciones según un criterio.

Relacionados con el procesamiento analítico

- R19. Visualizar cubos multidimensionales de datos relacionados con los casos registrados desde diferentes perspectivas de análisis.
- R20. Reorientar los cubos multidimensionales de datos.
- R21. Aumentar nivel de agregación en los datos de los cubos multidimensionales.



R22. Disminuir nivel de agregación en los datos de los cubos multidimensionales.

R23. Realizar el proceso ETL (extracción – transformación – carga).

R24. Visualizar reportes de pacientes registrados por día.

R25. Visualizar reportes de muestras recibidas por día.

R26. Visualizar reportes de determinaciones realizadas por día.

Relacionados con la administración, seguridad y protección

R27. Insertar datos de un usuario.

R28. Modificar datos de un usuario.

R29. Eliminar datos de un usuario.

R30. Visualizar datos de un usuario.

R31. Buscar datos de uno o varios usuarios según un criterio.

R32. Ordenar datos de los usuarios según un criterio.

R33. Insertar datos de un rol.

R34. Modificar datos de un rol.

R35. Eliminar datos de un rol.

R36. Visualizar datos de un rol.

R37. Buscar datos de uno o varios roles según un criterio.

R38. Ordenar datos de los roles según un criterio.

R39. Insertar datos de un acceso.

R40. Modificar datos de un acceso.

R41. Eliminar datos de un acceso.

R42. Visualizar datos de un acceso.

R43. Buscar datos de uno o varios accesos según un criterio.

R44. Ordenar datos de los accesos según un criterio.

R45. Insertar datos de un serotipo.



- R46. Modificar datos de un serotipo.
- R47. Eliminar datos de un serotipo.
- R48. Visualizar datos de un serotipo.
- R49. Buscar datos de uno o varios serotipos según un criterio.
- R50. Ordenar datos de los serotipos según un criterio.
- R51. Insertar datos de un área de salud.
- R52. Modificar datos de un área de salud.
- R53. Eliminar datos de un área de salud.
- R54. Visualizar datos de un área de salud.
- R55. Buscar datos de una o varias áreas de salud según un criterio.
- R56. Ordenar datos de las áreas de salud según un criterio.
- R57. Insertar datos de un municipio.
- R58. Modificar datos de un municipio.
- R59. Eliminar datos de un municipio.
- R60. Visualizar datos de un municipio.
- R61. Buscar datos de uno o varios municipios según un criterio.
- R62. Ordenar datos de los municipios según un criterio.
- R63. Insertar datos de una provincia.
- R64. Modificar datos de una provincia.
- R65. Eliminar datos de una provincia.
- R66. Visualizar datos de una provincia.
- R67. Buscar datos de una o varias provincias según un criterio.
- R68. Ordenar datos de las provincias según un criterio.
- R69. Iniciar sesión.
- R70. Terminar sesión.
- R71. Validar usuario.



Por otra parte, los requerimientos o requisitos no funcionales del sistema se refieren a las características que debe tener el mismo, que no se relacionan directamente con las respuestas que tiene que brindar al usuario [Jacobson, 2000]. Además, limitan la actuación del sistema a un ambiente de trabajo donde se debe ejecutar de forma correcta.

Relacionados con la apariencia o interfaz externa

- El diseño visual debe atraer al usuario y mantenerlo interesado en su uso.
- Todos los procesos incluidos dentro del sistema deben tener semejanza a los usados por los trabajadores cotidianamente, para disminuir la resistencia al cambio por parte de los usuarios finales.
- Los colores del sistema deben ser verde, blanco, gris y/o azul; colores que identifican al sector de la Salud.

Relacionados con la usabilidad

- El sistema debe ser fácil de usar y entender; pues los usuarios finales no son especialistas en la rama de la Informática.
- El sistema debe estar disponible las veinticuatro horas del día, los siete días de la semana.
- El sistema podrá ser explotado solamente en aquellas computadoras donde sea instalado el software, y siempre utilizando los sistemas de seguridad necesarios.

Relacionados con el software

- Las computadoras que vayan a hacer uso del sistema deben tener instalada la Máquina Virtual de Java.
- La resolución de pantalla debe ser 1024 x 786 píxeles por pulgada.
- La computadora servidora debe contar con la Plataforma Pentaho Open Source Business Intelligence, y el sistema gestor de bases de datos MySQL.



- Las computadoras clientes deben tener instalado un navegador Web.
- El sistema deberá correr en plataforma Windows y Linux, pues aunque la infraestructura de software del sector se basa en Windows, existe una estrategia a mediano plazo para migrar a software libre.
- Se debe garantizar la comunicación entre los componentes transaccionales y analíticos.

Relacionados con el hardware

- Se debe contar con un servidor de datos y un servidor de aplicaciones Web.
- Se debe contar con conexiones dúplex, para el envío y recepción de datos entre los servidores y entre las terminales y los servidores, a 100 Mbps, comunicación entre nodos y áreas con cableado de par trenzado o tecnología superior.
- Las computadoras clientes deben ser Pentium 3 a 133 Mhz con 32 Mb de RAM como mínimo.
- Los requerimientos de hardware del servidor, donde se hospede el sistema, y donde se instale el MySQL, están condicionados por los requerimientos mínimos de las plataformas (RAM de 512 Mb, espacio en disco de 1 Gb).

Relacionados con la seguridad

- La gestión de la información es restringida en cualquiera de los dos componentes principales del sistema.
- Sólo el administrador podrá administrar los usuarios.
- Sólo los especialistas podrán administrar la información del sistema.
- Se debe garantizar la confiabilidad de la información que se trasmite.
- Sólo el administrador de la computadora servidora podrá tener acceso a los ficheros fuentes y a las bases de datos del sistema.



- El sistema deberá ser capaz de controlar los usuarios que acceden al mismo y los accesos a las informaciones para cada uno de los mismos, restringiendo los accesos a sus funcionalidades de acuerdo a los roles.

Relacionados con la confiabilidad

- La información almacenada en el sistema debe corresponderse con el resultado de los procesos que tienen lugar en el negocio.

Relacionados con las políticas

- El sistema no debe violar las políticas trazadas dentro de la organización, y debe seguir las restricciones de la misma [MINSAP, 2004].

2.2.2. Actores del sistema

Los actores del sistema son los usuarios finales y sistemas externos que interactúan con el sistema, brindando información o recibiendo algún beneficio del mismo [Jacobson, 2000]. En la Tabla II se muestran los actores del sistema y su justificación.

Tabla II. Actores del sistema

Actor del sistema	Justificación
Técnico Laboratorio SUMA	Representa a la persona encargada de registrar los resultados de cada determinación, responsable del componente transaccional del sistema.
Directivo CPHE – SC	Representa a la persona encargada de analizar e interpretar los consolidados de datos de vigilancia, responsable del componente analítico del sistema.
Administrador del Sistema	Representa a la persona encargada de la administración y configuración del sistema (<i>back – end</i>).
Usuario restringido	Representa a todo usuario del sistema que debe autenticarse para poder acceder a los diferentes



	componentes del sistema, en correspondencia con su rol.
Herramienta ETL	Representa al sistema encargado de realizar la extracción, transformación y carga de los datos, desde el subsistema transaccional hasta el subsistema analítico.

En la Figura 2.2 se puede observar la relación que existe entre cada uno de los actores del sistema identificados en la tabla anterior.

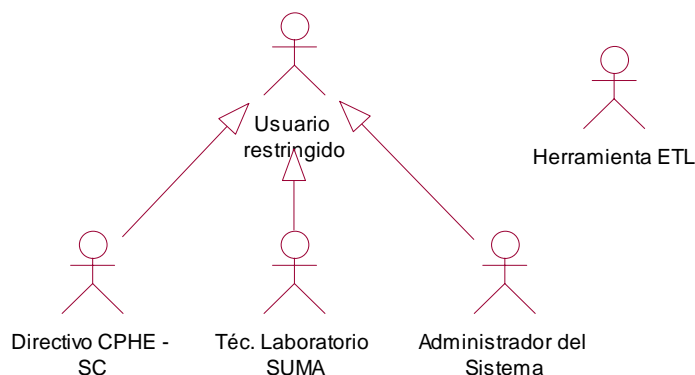


Figura 2.2. Actores del sistema

2.2.3. Diagrama de paquetes y casos de uso del sistema

Teniendo en cuenta los requerimientos funcionales identificados y los actores del sistema, se definieron cinco paquetes, subsistemas o lo que es lo mismo, módulos. Estos se relacionan en la Tabla III, con su justificación.

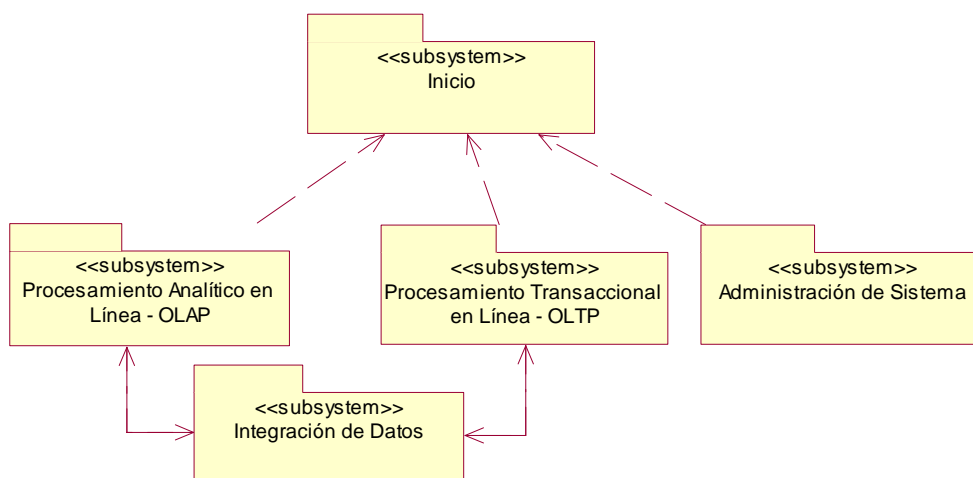
Tabla III. Módulos del sistema

Módulo	Justificación
Inicio	Este paquete agrupa todas las funcionalidades referidas a la autenticación del sistema.
OLTP Procesamiento	Este paquete agrupa todas las funcionalidades relacionadas con el procesamiento transaccional en



Transaccional en Línea	línea. Este paquete contiene la información básica y operacional del Sistema de Vigilancia Epidemiológica en cuestión; necesaria para poder poblar el componente analítico en línea de este sistema.
OLAP Procesamiento Analítico en Línea	Este paquete agrupa todas las funcionalidades que permiten el análisis en línea de la información del Sistema de Vigilancia Epidemiológica en cuestión.
Administración del Sistema	Contiene las funcionalidades identificadas para la administración y configuración del sistema.
Integración de Datos	Este paquete engloba las funcionalidades para la extracción de los datos operacionales, su transformación y carga en las estructuras multidimensionales del sistema (almacén de datos). Estas funcionalidades serán programadas haciendo uso de la herramienta <i>Pentaho Data Integration</i> .

En la Figura 2.3 se puede observar la relación que existe entre cada uno de los paquetes o subsistemas del sistema, identificados en la tabla anterior.



Fuente: Elaboración propia.

Figura 2.3. Diagrama de paquetes del sistema



El diagrama de caso de usos del sistema muestra gráficamente la relación existente entre los actores y procesos a informatizar en el sistema. Este diagrama está formado por casos de uso, actores, así como sus interacciones.

Este artefacto puede ser considerado como un acuerdo entre los desarrolladores y los clientes en cuanto a las funcionalidades que se van a incluir dentro del sistema a construir [Jacobson, 2000]. Estos fueron identificados teniendo en cuenta los requerimientos funcionales reflejados anteriormente. Los diagramas de casos de uso para cada paquete del sistema se pueden consultar en el Anexo 3.

2.3. Valoración de sostenibilidad de la solución propuesta

La sociedad aspira a un mundo mejor en los aspectos económico, social y ambiental. El concepto que expresa este deseo es el de desarrollo sostenible.

Constituye la valoración de sostenibilidad de un producto informático un proceso de evaluación de su impacto desde el diseño del proyecto, favoreciendo su autorregulación. Este proceso se lleva a cabo con un uso racional de recursos y la toma de decisiones adecuadas a las condiciones del contexto y el cliente [Concepción, 2006]. De ahí que se haya realizado una valoración de sostenibilidad del producto informático propuesto.

Dimensión administrativa

La confección de un producto informático, muchas veces, trae consigo un impacto administrativo. Este se refleja en los cambios que tienen lugar en la producción, los servicios y la economía de manera general en una organización.

Los beneficios que se obtendrán con este producto informático son positivos en gran medida, puesto que para la elaboración del mismo no es necesaria la adquisición de equipamiento, sino que se empleará el existente en el CPHE. Además no implica gasto alguno, pues las herramientas y tecnologías a utilizar son totalmente libre.

Por otra parte ayuda a mejorar los procesos de gestión de información del Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba, favoreciendo el



proceso de toma de decisiones, sin que su uso o implantación acarree gasto alguno. Para la explotación del sistema se cuenta con los requerimientos técnicos necesarios. Permite el ahorro de recursos materiales, humanos, tiempo y herramientas que son utilizadas para dicha gestión.

Dimensión socio – humanista

Cuando se hace alusión al impacto social, se hace referencia a cualquier alteración o cambio que se produce en las diferentes dimensiones de la realidad social del entorno o área de influencia del proyecto a desarrollar.

Con el producto informático se logra que el personal del CPHE de Santiago de Cuba que haga uso del mismo se sienta cómodo. Con este sistema se pretende que los usuarios trabajen con mayor eficiencia y rapidez, garantizando así mayor calidad en la gestión de información que genera el Sistema de Vigilancia Epidemiológica del Dengue en el territorio, que tribute a un mejor proceso de apoyo a la toma de decisiones en general.

Además, se obtiene un producto informático para el apoyo a la toma de decisiones, como resultado de la búsqueda de nuevos conocimientos en esta área, desde la implementación y uso de sistemas transaccionales y analíticos en línea. Esto hace que pueda ser generalizado en contextos similares.

El sistema mejora la calidad de vida de la sociedad, disminuyendo la contaminación y el daño al medio ambiente, perfeccionando la cultura profesional de los gestores de información y directivos, obligándolos a superarse en esta rama. El sistema no cierra ni genera empleos, sino que de manera directa favorece el proceso de toma de decisiones en el centro.

Dimensión ambiental

Se puede definir como cualquier alteración que se produzca en el medio ambiente al realizarse un proyecto o cualquier actividad humana.

Al favorecer la toma de decisiones, permitirá el ahorro de recursos que generan daños al medio ambiente, como papel y tinta de impresoras. Por otra parte,



permite un control más certero sobre la actividad administrativa del CPHE y específicamente sobre el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba. En este sentido, el sistema contribuirá a mejorar la calidad de vida de la población, y con ello el medio que les rodea, teniendo un impacto favorable en el medio ambiente.

No provoca daños físicos. No usa colores agresivos a la vista, ni genera contaminación por ruido. Favorece una interfaz agradable al entorno de usuario.

Dimensión tecnológica

Para su implantación, se prevé que los usuarios que hagan uso del sistema tengan conocimientos mínimos de toma de decisiones y vigilancia epidemiológica.

La característica de software libre hace que esté en constante perfeccionamiento por parte de sus desarrolladores, a través de nuevas plataformas de inteligencia de negocio. En tal sentido, se garantiza el perfeccionamiento del mismo, su adaptabilidad y flexibilidad ante nuevos cambios, así como mejores prestaciones.

De forma previsible se puede constatar que el producto informático a elaborar es perdurable en el tiempo, según la coherencia entre la necesidad social o problema que generó su elaboración y los recursos empleados para solucionarlo. Este proceso de valoración de sostenibilidad se realiza durante todo el ciclo de vida del software, de manera que se pueda corroborar la valoración realizada hasta el momento.

2.4. Construcción del sistema

Antes de acometer la tarea de construcción del sistema, se realizó un estudio de factibilidad. En tal sentido se estimó cuánto costaría, cuánto demoraría, y cuál sería la fuerza de trabajo necesaria para poder cumplir en tiempo con el desarrollo del software. Para ello se utilizó el Modelo Constructivo de Costos (COCOMO, por sus siglas en inglés), para estimar los datos correspondientes al software a desarrollar.



Específicamente, se utilizó la versión 2 del COCOMO, que se basa en puntos de función para realizar las estimaciones de las características del software [Boehm, 2000].

Después de realizar los cálculos pertinentes se obtuvo que el esfuerzo sería de 15.6 hombres/mes, lo que indica que si se desea realizar el sistema en un mes se necesitaría un equipo de 16 personas trabajando directamente en él.

En el caso del tiempo de desarrollo, se obtuvo que este fuera de 8 meses, en el que trabajarían dos personas directamente en la elaboración del sistema.

Cuando se analiza el costo del proyecto, se consideró el salario de los desarrolladores durante los 8 meses estimados. De esta forma se pudo determinar el costo del Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue, que fue de \$ 9,240.40 MN (aproximadamente \$ 462.03 USD).

Si se compara el costo que representa la atención de un solo paciente con dengue⁴ (aproximadamente \$ 594.00 USD) [Rodríguez, 2012], con el costo del proyecto; y además se tiene en cuenta los beneficios que traería el hecho de implantar un sistema como el propuesto, se considera que es factible económicamente el diseño e implementación del mismo.

Este sistema permitirá el ahorro de forma indirecta por concepto de combustible, insumos de laboratorio, entre otros. Además facilita una mayor y mejor gestión administrativa desde el orden epidemiológico de la enfermedad del dengue en Santiago de Cuba.

A continuación se tratan aspectos de las etapas de diseño e implementación del sistema, para dar solución a la situación problemática planteada. Se muestran artefactos que propone RUP para la descripción de las funcionalidades de la aplicación, el diseño de datos y se describe fundamentalmente la creación de las estructuras necesarias desarrolladas para el análisis OLAP.

⁴ Atención tanto hospitalaria como de vigilancia.



2.4.1. Diagrama de clases

En los diagramas de clases del diseño se muestra de forma gráfica las clases, atributos y operaciones que se definen para el sistema [Jacobson, 2000].

En el módulo de inicio se definió una clase servidora que contiene la información necesaria para establecer las conexiones a las bases de datos y la aplicación Web. Esta clase construye la página cliente con su formulario, donde el usuario podrá autenticarse.

Para el módulo de procesamiento transaccional se definió una clase servidora, que contiene la información necesaria para establecer las conexiones a las bases de datos y la aplicación Web. Se crearon además tres páginas clientes con sus respectivos formularios para la gestión de los datos relacionados con los pacientes registrados, las muestras tomadas y las determinaciones realizadas.

Para el módulo de procesamiento analítico se diseñaron dos clases servidoras, que contienen la información necesaria para establecer las conexiones a las bases de datos y ejecutar los ficheros que contienen la definición XML de los cubos multidimensionales y los reportes respectivamente. Estas clases servidoras construyen dos páginas clientes con sus formularios, que permiten la gestión de los cubos y los reportes relacionados con los casos de dengue desde diferentes perspectivas de análisis.

En el caso del módulo de administración y seguridad, fue identificada una clase servidora que contiene la información necesaria para establecer las conexiones a las bases de datos y la aplicación Web. Esta clase servidora construye ocho páginas clientes con sus respectivos formularios, que en su conjunto permiten la gestión de los datos relacionados con los usuarios, los roles, los accesos al sistema, así como los elementos utilizados en el módulo de procesamiento transaccional (áreas de salud, municipios, provincias, serotipos, etc.).

Una muestra representativa de los diagramas de clases que ilustran la estructura de la solución propuesta puede ser consultada en el Anexo 4.



2.4.2. Diagrama de clases persistentes y modelos de datos

Las clases persistentes representan la información que tiene una vida duradera dentro del sistema informático [Jacobson, 2000]. Dichas clases deben tener un tratamiento especial, pues van a definir directamente la composición final de la base de datos del sistema informático.

Luego de analizada la información persistente del sistema, y las características propias del gestor de base de datos, se obtuvo un diseño de la base de datos que garantiza los mecanismos de persistencia y recuperación de la información.

La siguiente figura muestra el diagrama de clases persistentes para el componente transaccional y de administración del sistema, el cual se encuentra normalizado hasta la Forma Normal de *Boyce – Codd*.

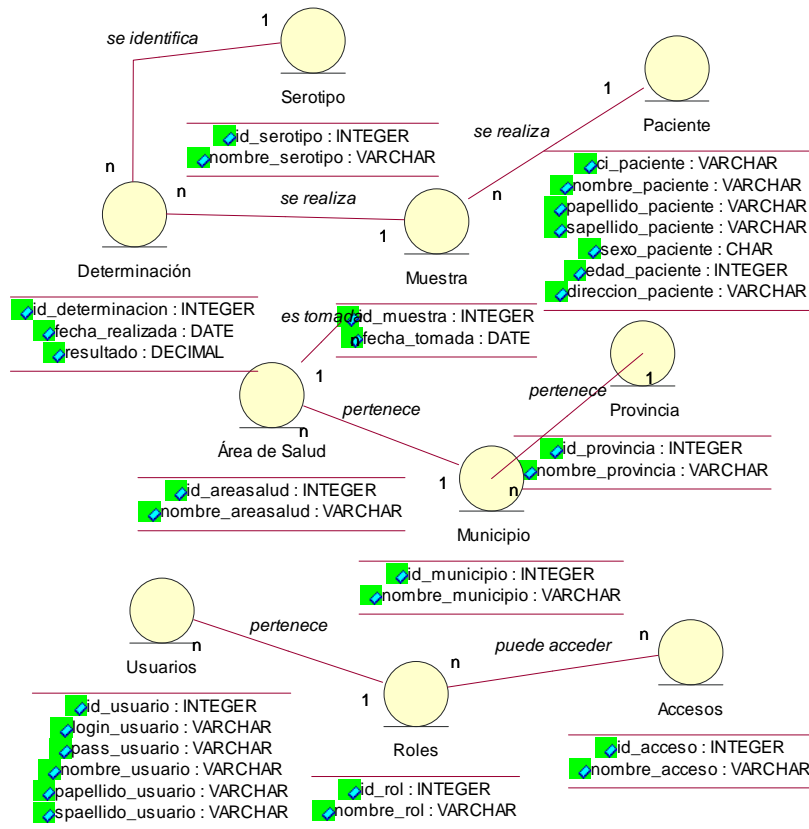


Figura 2.4. Diagrama de clases persistentes del subsistema OLTP y Administración



En el caso del diagrama de clases persistentes para el componente analítico del sistema, las estructuras fueron identificadas a partir de la aplicación de la metodología HEFESTO, propuesta por Ricardo D. Bernabeu [Bernabeu, 2009].

En tal sentido, se definieron los indicadores y las perspectivas de análisis, resultando los siguientes:

- Total de casos realizados⁵ por lugar y para una fecha determinada.
- Total de casos confirmados⁶, por serotipo, por lugar y para una fecha determinada.
- Total de casos realizados y confirmados por lugar y para una fecha determinada.

A partir de las necesidades de análisis anteriores, se identificaron como medidas (indicadores) el total de casos realizados y el total de casos confirmados. Las dimensiones (perspectivas de análisis) definidas fueron: lugar, fecha y serotipo.

El agregador a utilizar para cada una de las medidas es la función de sumariación (SUM), ya que lo que se busca en el análisis es el total en la intersección de cada una de las dimensiones.

Luego se procedió a elaborar el modelo conceptual. Básicamente se definió la relación entre cada una de las dimensiones con las medidas necesarias en cada análisis, y se determinó la correspondencia entre este y el diagrama entidad – relación (DER) del componente OLTP. Se definió además la granularidad con que se analizaban los datos, decidiendo finalmente implementar un esquema “en estrella”, a partir de las siguientes estructuras:

- *Casos*: tabla que contendrá los hechos: casos presuntivos o realizados, y casos confirmados.
- *Serotipo*: tabla de dimensión que contendrá los diferentes serotipos de dengue desde los cuales se podrá realizar el análisis.

⁵ Este indicador se define en las reglas del negocio (ver Anexo 1).

⁶ Ídem al anterior.



- *Lugar*: tabla de dimensión que contendrá los diferentes escenarios de salud donde se detecte la presencia de la enfermedad. Para esta tabla se definieron los campos: área de salud, municipio, provincia.
- *Tiempo*: tabla de dimensión que contendrá el registro de las fechas en que se registren los casos presuntivos y confirmados. Para esta tabla se definieron los atributos: año, trimestre, mes y día.

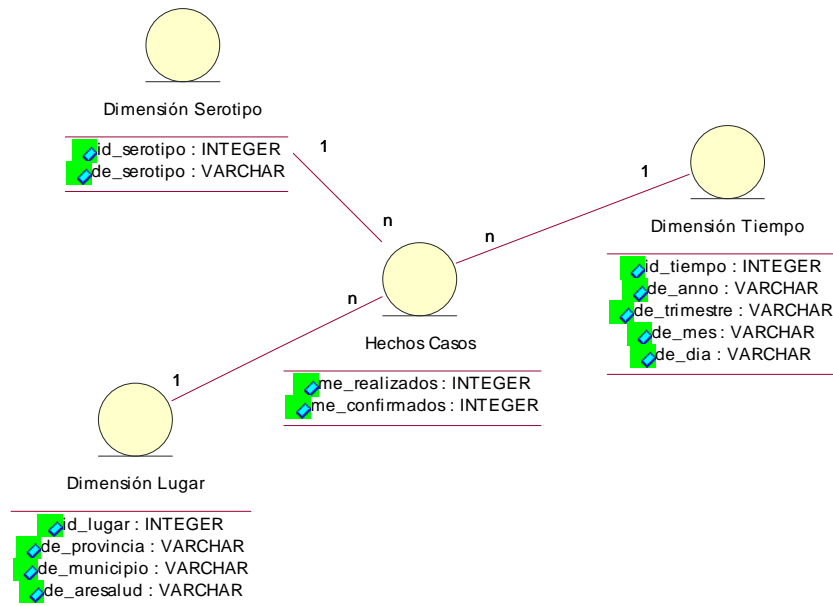


Figura 2.5. Diagrama de clases persistentes para el subsistema OLAP

Como se puede apreciar, las tablas de las dimensiones están desnormalizadas, una característica fundamental para este esquema, ya que es la vía para poder “navegar” por las dimensiones en dependencia del nivel de granularidad deseado.

Se generaron dos modelos de datos para así representarlos físicamente en las dos bases de datos concebidas para el sistema: bd_dengue_oltp y bd_dengue_olap. Estos modelos se muestran en el Anexo 5.

2.4.3. Diseño visual

El diseño visual del Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba se ha hecho de tal



forma que garantice una correcta interacción con los usuarios. Estos deben extraer la mayor cantidad de información en un período de tiempo breve.

La interacción usuario – sistema se realiza cuando el técnico de Laboratorio SUMA inserta o actualiza los datos relacionados con los pacientes, las muestras y las determinaciones; cuando el administrador gestiona cada uno de los módulos del sistema; así como cuando los usuarios, o cuando los directivos del CPHE analizan el volumen de información generado por el OLTP, para la toma de decisiones desde el orden epidemiológico de esta enfermedad.

Los colores predominantes son claros, para evitar el cansancio de los usuarios finales durante la utilización del sistema. De la misma manera se utilizaron textos en color negro para resaltar la información.

Las clases de presentación de información se diseñaron de forma tal que tuvieran la información mínima necesaria para facilitar la extracción de la información que pudiera tener utilidad. En el Anexo 6 se pueden observar algunas pantallas típicas del sistema.

2.4.4. Soporte para el análisis OLAP en el sistema

En el Epígrafe 2.4.2 se reflejó la concepción del modelo de datos que soporta la base de datos multidimensional del sistema. Para poblar dichas estructuras desde la base de datos operacional del mismo, se utilizó el *Pentaho Data Integration* como herramienta encargada para llevar a cabo el proceso de extracción, transformación y carga de los datos.

Para ello se definió un *job* que consta de dos pasos: paso de inicio, en el que se define el tipo de *schedule* o cronograma de ejecución, y una sola transformación, en la que se lleva a cabo el ETL de todos los datos. Este *job* fue incluido en la herramienta que provee el servidor de Pentaho para ejecutar tareas, y su ejecución estará condicionada por la frecuencia que se le programe en dicha herramienta. En la Figura 2.6 se muestra el diagrama que define a este *job*.



Figura 2.6. Diagrama que define el job actualizar_dw_dengue

Para la transformación concebida, que cuenta con doce pasos, se definieron las conexiones con cada una de las dos bases de datos. En la Figura 2.7 se muestra el diagrama que define esta transformación.

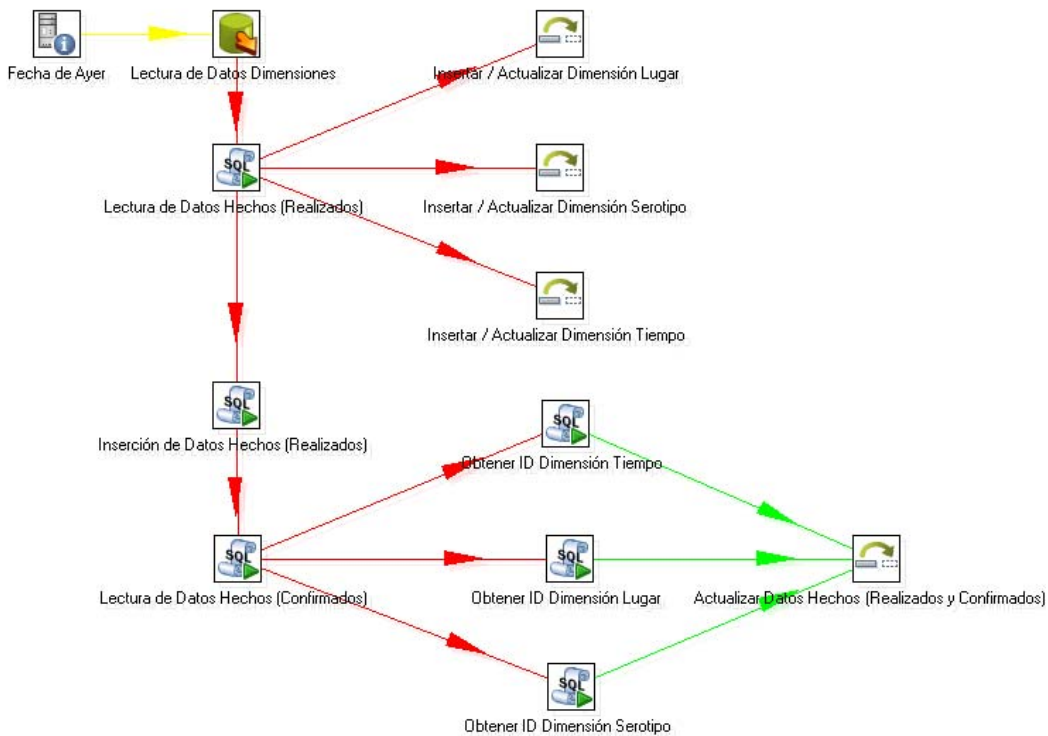


Figura 2.7. Diagrama que define la transformación actualizar_dimensiones

Inicialmente se toma la fecha del día anterior. Con esta información se leen los datos de las dimensiones y hechos desde la base de datos operacional, para los cuales fueron insertados datos en la fecha que se pasa como parámetro.

Con estos datos, se ejecutan tres pasos de inserción/actualización en las dimensiones, y un paso de ejecución de scripts SQL en la tabla de hechos con valores de medidas por defecto. Se ejecutan dos pasos de ejecución de scripts



SQL, para obtener las medidas para cada combinación de dichas dimensiones desde la base de datos operacional. Por último, se recuperan las llaves primarias de cada dimensión y son actualizadas las medidas en la tabla de hechos.

Una vez garantizada la integración de las bases de datos del sistema, se diseñó un esquema XML en el que se definen los tres cubos multidimensionales creados para el análisis, solicitados por el cliente. Para esto se utilizó la herramienta *Pentaho Schema Workbench*, la que posibilitó la creación y publicación de dichos cubos en la plataforma Pentaho.

El primer cubo creado (casos realizados) consta de dos dimensiones: lugar y fecha. En la Figura 2.8 se muestra gráficamente la construcción de este cubo.

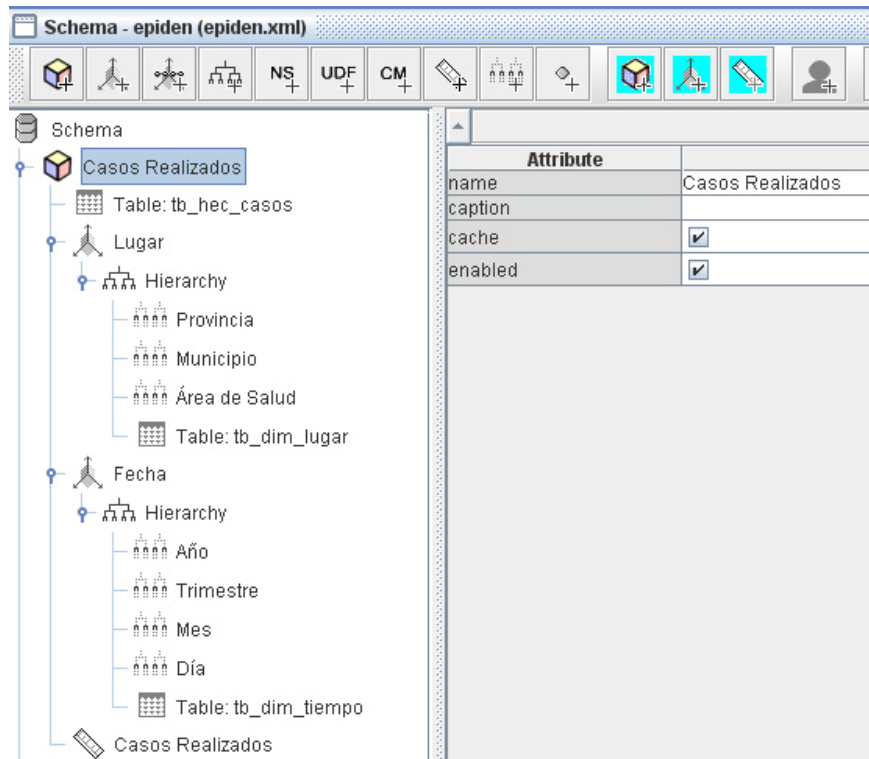


Figura 2.8. Diseño visual de uno de los cubos en el Schema Workbench

Como se pudo apreciar, en esta herramienta se definió la tabla de hechos desde la cual el cubo recuperaría los datos (tb_hec_casos).



En el caso de la dimensión *lugar* para este cubo, se le definió la tabla de la dimensión desde la cual tomaría los datos (*tb_dim_lugar*), y el campo de llave extranjera en la tabla de hechos que permite relacionarla con dicha tabla. Se le incluyó además una jerarquía con tres niveles, que se corresponden con los tres campos no llave de esta dimensión (área de salud, municipio y provincia).

En el caso de la dimensión tiempo para este cubo, se le definió la tabla de la dimensión desde la cual tomaría los datos (*tb_dim_tiempo*), y el campo de llave extranjera en la tabla de hechos que permite relacionarla con dicha tabla. Se le incluyó además una jerarquía con cuatro niveles, que se corresponden con los cuatro campos no llave de esta dimensión (día, mes, trimestre y año).

Se definió además la medida que se mostraría en el cubo para las dimensiones antes descritas, que se corresponde con el atributo de casos presuntivos (*me_realizados*). Esta medida se muestra en el cubo utilizando el *aggregator* SUM.

En el Anexo 7 se muestra el código XML que define el esquema Mondrian que contiene el cubo descrito anteriormente, así como los otros dos cubos diseñados y publicados en la plataforma.

La inclusión de estos cubos en el sistema se llevó a cabo a partir de la creación de una secuencia de acción (*xaction*) para cada cubo. Esto se logró con la herramienta *Pentaho Design Studio*. Una vez creadas, estas fueron publicadas en el servidor Pentaho, y sus URL fueron incluidas en el sistema. Estos son vistos utilizando el *JPivot Viewer* de la plataforma, para visualizar las vistas de análisis. Desde estas vistas, el usuario podrá generar reportes en varios formatos, navegar por las dimensiones creadas, aumentar o disminuir el nivel de detalle para el análisis, rotar los ejes del cubo, entre otras funcionalidades.

En la Figura 2.9 se muestra una vista de análisis desde el sistema para uno de los cubos creados.



Lugar	Fecha	Medidas	
		● Casos Realizados	● Casos Confirmados
Lugares	Fechas	58	30
	2012	58	30
	1er Trimestre	22	14
	2do Trimestre	36	16
Santiago de Cuba	Fechas	58	30
	2012	58	30
	1er Trimestre	22	14
	2do Trimestre	36	16
Santiago de Cuba	Fechas	58	30
	2012	58	30
	1er Trimestre	22	14
	2do Trimestre	36	16
28 de Septiembre	Fechas	1	0
	2012	1	0
	2do Trimestre	1	0
30 de Noviembre	Fechas	24	14
	Fechas	24	14

Figura 2.9. Vista de análisis de uno de los cubos creados y publicados

En el caso de los reportes, se crearon los tres identificados como requerimientos funcionales. En la Figura 2.10 se muestra el ejemplo de un reporte generado.



LABORATORIO SUMA. CPHE - SANTIAGO DE CUBA

Reporte Diario de Determinaciones Realizadas

Determinaciones realizadas el día: 19.04.2012

Fecha de emisión de este reporte: 22.04.2012

Cód. Muestra	Cód. Determinación	Serotipo	Fecha Realizada	Resultado
1	6	Serotipo 6	19 / 04 / 2012	455,00
1	7	Serotipo 6	19 / 04 / 2012	11,00
2	5	Serotipo 5	19 / 04 / 2012	13,96

Total de Determinaciones: 3

Figura 2.10. Muestra de uno de los reportes creados, publicados y generados

Los reportes fueron diseñados utilizando la herramienta *Pentaho Report Designer*, que posibilitó su diseño y publicación en el servidor. Para cada uno se utiliza el



parámetro: fecha de reporte, con el cual el reporte es generado y mostrado al usuario. Además se brinda la opción de seleccionar el formato en que se desea ver el reporte, entre los que se destacan: HTML, PDF, entre otros.

2.4.5. Tratamiento de errores y seguridad

En el Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba se disminuye la cantidad de información que el usuario debe introducir por medio del teclado, ya que gran parte de esta es administrada desde el mismo sistema. Esto hace que disminuye considerablemente la cantidad de errores que puedan ser introducidos por esta vía.

La información introducida es validada constantemente y el usuario es retroalimentado en caso de que se detecte algún error.

Para la programación de los subsistemas de OLTP y Administración del Sistema, se utilizó el IDE: PHP Generator for MySQL, y con este, el lenguaje de programación PHP. Este lenguaje brinda mecanismos de tratamiento de excepciones para la manipulación de los errores, empleándose fundamentalmente para el acceso a datos.

La política de tratamiento de errores empleada disminuye considerablemente la cantidad de problemas que puedan ser introducidos en este sistema.

En el caso de la seguridad, se debe señalar que el sistema en su totalidad es restringido, no tiene accesos públicos, por lo que para acceder al mismo es necesario tener un usuario y una contraseña válidos. Lo primero que un usuario debe hacer para utilizar el sistema es autenticarse.

Es importante señalar que el administrador sólo puede añadir o modificar usuarios, roles y accesos. Los usuarios que se autentican como parte del rol de administrador no pueden actualizar o modificar información relacionada con el Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba.



Los usuarios que entran al sistema y que tienen el rol de especialista son dirigidos a las funcionalidades que permiten actualizar o modificar la información relacionada con este Sistema.

2.4.6. Despliegue del sistema

El diagrama de despliegue es el artefacto que muestra como se distribuyen los componentes del sistema físicamente por cada uno de los equipos denominados nodos. Entre estos nodos se definen relaciones que indican la interacción que existe entre ellos. Dichas interacciones se realizan utilizando medios de comunicación [Jacobson, 2000].

Se definió un servidor central, en el que se encuentra la plataforma Pentaho Open Source Business Intelligence, con el sistema, que ejecuta todas las funcionalidades que sobre ella se llevan a cabo. Este nodo contiene también las dos bases de datos que utiliza la aplicación: bd_dengue_oltp (relacional) y bd_dengue_olap (multidimensional). Se definió además un nodo cliente que representa las computadoras clientes, desde las cuales los usuarios interactuarán con el sistema.

En la siguiente figura se muestra cómo va a desplegarse el sistema.

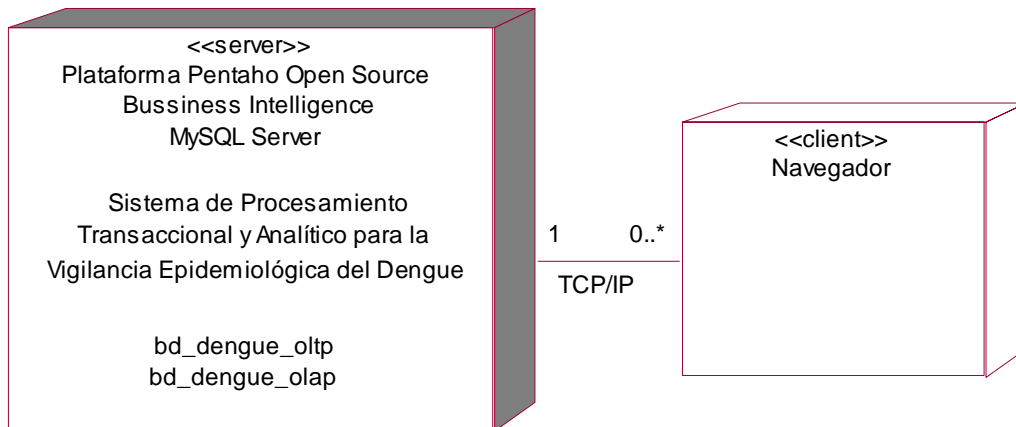


Figura 2.11. Diagrama de despliegue del sistema



2.4.7. Estructura general del sistema

Para comenzar a utilizar el Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica de Dengue se necesita tener las instalaciones de los siguientes software: MySQL 5.0 para el servidor de bases de datos, un servidor Apache para el servidor Web, y la plataforma Pentaho Open Source Business Intelligence para el despliegue de las herramientas integradas a ésta. Luego se deben restaurar las bases de datos (operacional y multidimensional) con un *backup* inicial que contiene información sobre los esquemas físicos de estas, así como ficheros normativos, indispensables para el correcto funcionamiento de la aplicación.

Por otra parte, se deben restaurar también las bases de datos que utiliza Pentaho para su funcionamiento, mantenimiento y administración (*hibernate*, *quartz*, y *sampledata*); así como también se deberán publicar con las herramientas correspondientes cada uno de los elementos de análisis, ETL y reportes.

Módulo Inicio

Lo primero que un usuario debe hacer para utilizar el sistema, es autenticarse desde el área de autenticación de la Página Inicio.

Módulo de Procesamiento Transaccional en Línea

En este módulo se podrán gestionar los datos relacionados con la siguiente categoría:

- *Actividades de laboratorio*: pacientes registrados, muestras recibidas, y determinaciones realizadas.

Siempre y para cada categoría, se mostrarán en un *datagrid* los elementos almacenados por el sistema, y las funcionalidades básicas de insertar, editar, y eliminar. Estas operaciones de almacenamiento y recuperación de información se llevan a cabo sobre la base de datos operacional: *bd_dengue_oltp*.

Además se implementaron las funcionalidades de ordenación de registros de forma ascendente o descendente, así como la de búsqueda (filtro) para todos los



campos o por un campo específico. Esto flexibiliza y agiliza aún más el proceso de gestión de estos datos.

Módulo de Procesamiento Analítico en Línea

Para el usuario redireccionado al módulo de procesamiento analítico en línea, las funcionalidades se muestran agrupadas en dos categorías:

- *Análisis de la información:* casos realizados por fecha y lugar; casos confirmados por serotipo, fecha y lugar; y casos realizados y confirmados por serotipo, fecha y lugar.
- *Reportes de actividad:* reporte de pacientes registrados por día, reporte de muestras recibidas por día, y reporte de determinaciones realizadas por día.

Bajo la categoría de análisis de la información se tiene acceso a cada uno de los tres cubos creados para el análisis en línea, los cuales son llamados desde la aplicación, pero presentados con el *JPivot Viewer*, que se ejecuta sobre la plataforma Pentaho. Desde esta categoría, el usuario podrá analizar la información desde las perspectivas de análisis que desee o que necesite a la hora de tomar una decisión. De esta forma podrá aumentar o disminuir el nivel de detalle de los datos, reorientar el cubo, entre otras funcionalidades. Las operaciones de recuperación de información se lleva a cabo sobre la base de datos multidimensional: `bd_dengue_olap`.

Bajo la categoría de reportes, el usuario podrá visualizar los reportes correspondientes, pudiendo configurar los parámetros deseados para el reporte, relacionados con el día de reporte y el formato de visualización de los mismos.

Módulo de Administración y Seguridad

Si el usuario es redireccionado al módulo de administración y seguridad, el usuario podrá gestionar datos relacionados con las siguientes categorías:

- *Seguridad:* usuarios del sistema, roles, accesos, y accesos por roles.
- *Módulo OLTP:* áreas de salud, municipios, provincias, y serotipos.



Al igual que para el módulo de procesamiento transaccional en línea, siempre y para cada categoría, se mostrarán en un *datagrid* los elementos almacenados por el sistema, y las funcionalidades básicas de insertar, editar, y eliminar. Estas operaciones de almacenamiento y recuperación de información se lleva a cabo sobre la base de datos operacional: *bd_dengue_oltp*.

También se implementaron las funcionalidades de ordenación de registros de forma ascendente o descendente, así como la de búsqueda (filtro) para todos los campos o por un campo específico, lo que flexibiliza y agiliza aún más el proceso de gestión de estos datos.

Módulo de Integración de los Datos

El funcionamiento de este módulo fue descrito previamente en el Subepígrafe 2.4.4. Las operaciones de recuperación y almacenamiento de la información se llevan a cabo sobre ambas bases de datos del sistema: *bd_dengue_oltp* y *bd_dengue_olap*.

En el Anexo 6 se pueden observar algunas pantallas típicas del sistema.

2.5. Utilización del sistema para el apoyo a la toma de decisiones

El sistema elaborado permite el registro y control de la información relacionada con el Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba. De esta forma refleja datos de cada uno de los casos sometidos a las pruebas de laboratorio SUMA, con lo que se garantiza la centralización y con ello la integridad de la información. En este sentido, se reduce la cantidad de errores cometidos en la emisión de partes, y se garantiza la disponibilidad de la información en todo momento, a través de las diferentes estructuras de almacenamiento de información concebidas en el sistema.

Por otra parte, se implementaron técnicas y herramientas para el análisis OLAP que permiten en su conjunto la recuperación, evaluación y presentación de la información al usuario de forma dinámica. Se previó además la navegación por cada una de las perspectivas de análisis, en la que el usuario pudiera aumentar o



disminuir el nivel de detalle de la información en dependencia de las necesidades de información que precise, en este caso el decisor para la toma de decisiones.

Para ilustrar lo anterior, por ejemplo, a los efectos del análisis OLAP para el apoyo a la toma de decisiones con el sistema; con el cubo creado: “*casos realizados y casos confirmados por fecha y lugar*”, el usuario podría responder las siguientes preguntas de análisis:

- ¿Cuántos casos se han reportado para un período determinado, en un lugar determinado?
- ¿Cómo se han comportado los casos confirmados para un período determinado, en un lugar determinado?
- ¿Cómo se han comportado los casos realizados y/o confirmados durante varios períodos de tiempo, en un lugar determinado?
- ¿En qué período del año/trimestre/mes se percibe históricamente un aumento o una disminución de casos confirmados?
- ¿Cómo se ha comportado la relación casos confirmados/casos realizados para un período de tiempo determinado?
- ¿Qué lugar(es) presenta(n) mayor número de casos confirmados, durante un período de tiempo determinado?

Por otra parte, permitiría evaluar otros aspectos, entre ellos:

- ¿Cuán efectivo es el método clínico para el diagnóstico de la enfermedad?
- ¿Cuán efectivo es el método epidemiológico para el diagnóstico de la enfermedad?
- ¿Pueden asociarse las tasas de morbilidad con algún elemento particular de las perspectivas de análisis, por ejemplo: un período determinado o un lugar determinado, a un nivel de detalle determinado?

Las respuestas a estas interrogantes posibilitarían tomar decisiones en torno a:



- La determinación de fases epidemiológicas de la enfermedad por áreas de salud, municipios, provincias, etc.
- La determinación de períodos que históricamente tienen asociados un aumento o una disminución de casos.
- La determinación de la relación existente entre casos realizados y casos confirmados, relación que en cierta medida permitiría valorar la aplicación del método clínico y/o del método epidemiológico empleado por el personal facultativo para diagnosticar la enfermedad.
- La determinación de lugares a diferentes niveles de detalles, que presentan mayores tasas de morbilidad asociadas a esta enfermedad por períodos.
- La asociación de la variabilidad de las tasas de morbilidad asociadas a esta enfermedad con períodos o lugares determinados, o la combinación de estos.
- La planificación de recursos, insumos y materiales tanto para la labor preventiva (lucha contra el vector que produce la enfermedad), como para la ejecución del procedimiento biotecnológico de diagnóstico, por períodos y lugares priorizados.

Como se puede observar, con la utilización de este cubo se puede analizar e interpretar el comportamiento y tendencia de la enfermedad del dengue según las diferentes perspectivas de análisis. Muchos de estos aspectos era casi imposible evaluarlos o simplemente no se evaluaban, por la complejidad del análisis en dependencia del nivel de detalle que se necesitara. Además, aun evaluando estos aspectos con consultas SQL tradicionales, a muchos de ellos no se les podría dar respuesta. De ahí la importancia y pertinencia del sistema elaborado.

2.6. Valoración del sistema utilizando criterio de expertos

Fue necesario conocer si el Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba apoya el



proceso de decisiones en el CPHE. Para ello se utilizó el Método Delphi, el cual se basa en las valoraciones individuales de expertos y usuarios [Cruz, 2009].

Mediante este método se elimina el peligro concreto que representan los individuos líderes, como silenciadores de la creatividad individual del resto de los expertos o usuarios consultados.

Se aplicó una primera encuesta para determinar el coeficiente de competencia de los expertos seleccionados en el tema: sistema de apoyo a la toma de decisiones durante la vigilancia epidemiológica del dengue (ver Anexo 8). Fueron consultados 32 profesionales. De estos se seleccionaron los expertos en el tema de referencia, con coeficientes de competencia alto (21) y medio (4).

En la caracterización de los 25 expertos seleccionados se destaca el nivel de conocimientos que poseen sobre vigilancia epidemiológica del dengue, y los sistemas de apoyo a la toma de decisiones, así como conocimientos medios y/o avanzados de sistemas informáticos. Los seleccionados tienen una alta preparación académica, todos son graduados de nivel superior, en su mayoría epidemiólogos, licenciados en Ciencias de la Computación, ingenieros en Ciencias Informáticas y licenciados en Tecnología de la Salud. El promedio de años de experiencia laboral es de 11 años.

De ellos, el 12% está formado por doctores en Ciencias de la Salud; el 76%, por másteres en Salud Pública, Enfermedades Infecciosas, Informática Médica, y especialistas de primer y segundo grado en Higiene y Epidemiología, para un total de 88%. El resto son investigadores en general, conocedores también de temas de administración. A todos se les aplicó la encuesta para arribar a conclusiones sobre el sistema.

La segunda encuesta permitió determinar el grado de importancia de las características del Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba como parte de un sistema de apoyo a la toma de decisiones (ver Anexo 9). Previamente fue sometida a la consideración de los expertos seleccionados, que opinaron que los



aspectos contenidos eran pertinentes para evaluar y valorar el sistema. Finalmente la encuesta incluyó 15 indicadores.

Además en esa encuesta se incluyeron preguntas abiertas en las que los expertos podían incluir sus valoraciones y opiniones personales del tema investigado. Estas preguntas constituyeron un importante punto de partida para el mejoramiento del sistema.

Una vez realizados los cálculos pertinentes, se determinó que el 33% de los aspectos incluidos en la encuesta los expertos lo valoran de muy relevante; el 47%, de bastante relevante y el 20%, de relevante. Los expertos no reflejaron ningún aspecto como no relevante o poco relevante (ver Anexo 10).

Los aspectos señalados de muy relevantes a juicio de los expertos fueron:

- Flexibilidad y mantenimiento del sistema.
- Integración de los datos entre las bases de datos del sistema.
- Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del dengue.
- Facilidades para la manipulación de los cubos multidimensionales.
- Generación de reportes parametrizados.

Luego de analizar el resultado de la encuesta aplicada, y teniendo en cuenta las sugerencias reflejadas, se puede decir que los expertos realizaron una valoración favorable del sistema. Por otra parte consideran factibles las funcionalidades implementadas para la gestión y el apoyo al proceso de toma de decisiones en el territorio respecto a la enfermedad en cuestión.

2.7. Conclusiones del capítulo

En este capítulo se describen las fases de diseño e implementación del Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue en Santiago de Cuba. Este proceso se mostró a través de los artefactos que propone la metodología RUP. Se determinaron las características



fundamentales del mismo, qué personas y entidades tienen acceso a él y con qué finalidad. También fue posible determinar los niveles de seguridad necesarios para mantener la información confiable y segura.

Se detalló cómo el sistema soporta el análisis OLAP y cuáles técnicas y herramientas fueron creadas para ello. Se destaca la integración de las bases de datos del sistema, la creación, publicación y presentación de cubos multidimensionales y reportes a través de las herramientas de desarrollo integradas a la plataforma Pentaho Open Source Business Intelligence.

Se obtuvo una valoración favorable del sistema utilizando el criterio de experto según el Método Delphi. El sistema construido satisface los requerimientos identificados al inicio de su construcción.

De manera general el sistema garantiza la integridad, consistencia, confiabilidad y disponibilidad de la información. Estas características favorecen el proceso de gestión y de apoyo a la toma de decisiones durante la vigilancia epidemiológica del Dengue.



CONCLUSIONES

Con el desarrollo e implantación del Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue para la gestión de información y el apoyo al proceso de toma de decisiones, se cumplió con el objetivo propuesto en esta investigación.

Este software soporta la aplicación de técnicas de procesamiento transaccional y analítico en línea, enfocado a la explotación de grandes volúmenes de datos del Sistema de Vigilancia Epidemiológica del Dengue en Santiago de Cuba. Tiene como finalidad el análisis, interpretación y presentación de la información para la toma de decisiones en los distintos niveles del sector de Salud en la provincia.

El producto informático obtenido proporciona un conjunto de herramientas que permiten gestionar, analizar, interpretar y comparar tendencias desde diferentes perspectivas de análisis de esta enfermedad y su trasfondo epidemiológico en el territorio. El sistema, a través del uso de herramientas matemático - computacionales, permite que el proceso de toma de decisiones sea más dinámico y eficiente.



RECOMENDACIONES

Sobre la base de la investigación realizada y la experiencia obtenida durante su realización y para continuar el desarrollo de trabajos relacionados con los sistemas para la gestión de información de apoyo a la toma de decisiones en la vigilancia epidemiológica de enfermedades, se realizan las siguientes recomendaciones:

- Generalizar el sistema elaborado en otros centros de Higiene y Epidemiología de otras provincias, ya que los sistemas de vigilancia epidemiológica para cualquier enfermedad en el país, están regulados por normas y procedimientos establecidos dentro del Sistema Nacional de Salud.
- Dar continuidad a esta investigación a través de proyectos que permitan la generación de reportes de forma automática, así como la generación de otros cubos multidimensionales.
- Desarrollar aplicaciones similares para otras enfermedades que resulten de interés desde el punto de vista epidemiológico para los administradores del MINSAP a diferentes niveles.



REFERENCIAS BIBLIOGRÁFICAS

- [Andrés, 1999] Andrés, A. et al. *Vigila: un instrumento para la vigilancia epidemiológica de enfermedades transmisibles*. Rev Cubana Med Trop [online]. 1999, vol.51, n.1. ISSN 0375-0760. [Consultado: febrero de 2012].
- [Bernabeu, 2009] Bernabeu, R. D. *DATA WAREHOUSING: Investigación y Sistematización de Conceptos – HEFESTO: Metodología propia para la Construcción de un Data Warehouse*. Córdoba, Argentina. 2009.
- [Boehm, 2000] Boehm. *SW Cost Estimation with COCOMO II. Model Definition Manual*. Prentice Hall. 2000.
- [Buchanan, 2006] Buchanan, L. et al. *Breve historia de la Toma de Decisiones*. Harvard Deusto Business Review. No.148. 2006.
- [Cohen, 1996] Cohen, D. *Sistemas de información para la toma de decisiones*. Segunda Edición. McGRAW – HILL. New York. 1996.
- [Concepción, 2006] Concepción, R. *La gestión de proyectos Informáticos sostenibles*. Tesis en opción al título de máster en Gestión Ambiental. Universidad de Holguín. 2004.
- [Cruz, 2009] Cruz, M. *El método Delphi en las investigaciones educativas*. Editorial Academia. La Habana. 2009.
- [Date, 2003] Date, C. J. *Introducción a los Sistemas de Bases de Datos*. Editorial Félix Varela, La Habana. 2003.
- [Harkey, 1996] Harkey J., et al. *The essential client/server survival guide*. John Wiley & Sons. 1996.
- [Inmon, 2002] Inmon, W. H. *Building the Data Warehouse*. Third Edition. Wiley Computer Publishing. New York. 2002.
- [Jacobson, 2000] Jacobson, I. et al. *El Proceso Unificado de Desarrollo de Software*. Addison – Wesley. 2000.



- [Kimball, 2002] Kimball, R., et al. *The Data Warehouse Toolkit. Second Edition. The Complete Guide to Dimensional Modeling.* Wiley Computer Publishing. New York. 2002.
- [Kouri, 2011] Kouri, G. *El dengue, un problema creciente de salud en las Américas.* Rev Cubana Salud Pública [online]. 2011, vol.37, suppl.5. ISSN 0864-3466. [Consultado: febrero de 2012].
- [Larman, 2000] Larman, C. *Applying UML and Patterns. An Introduction to Object – Oriented Analysis and Design and the unified Process.* Second Edition. Prentice Hall. México. 2000.
- [Mateu, 2004] Mateu, C. *Desarrollo de Aplicaciones Web.* Fundació per a la Universitat Oberta de Catalunya. Barcelona. 2004.
- [McClure, 1978] McClure, C. R. *The information rich employee and information for decision making: Review and comments.* Information Processing and Management, 1978, vol. 14.
- [MINSAP, 1995] MINSAP. *Sistema de Vigilancia en Salud. Unidad de Análisis y Tendencias en Salud.* La Habana. 1995.
- [MINSAP, 2004] MINSAP. *PNO Dengue.* Laboratorio Provincial SUMA. Centro Provincial de Higiene, epidemiología y Microbiología. Santiago de Cuba. 2004.
- [Moss, 2003] Moss, L. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications.* 2003
- [Nader, 2006] Nader, J. *Sistema de Apoyo Gerencial Universitario.* Tesis de magister en Ingeniería del Software. Instituto Tecnológico de Buenos Aires. 2006.
- [Pence, 2008] Pence, N., Creeth, R. *An Introduction to OLAP.* Documento en línea. [Consultado: diciembre de 2008]. Disponible en: <http://www.OLAPreport.com/>.



- [Pentaho, 2008] Pentaho. Open Source Business Intelligence. *Mondrian Schema Workbench*. Documento en línea. 2008 [Consultado: julio de 2009]. Disponible en: <http://wiki.pentaho.com/display/PMOPEN/Mondrian+Schema+Workbench>.
- [Rodríguez, 1999] Rodríguez, R., et al. *El establecimiento de sistemas de información en servicios de atención de salud: Guía para el análisis de requisitos, especificación de las aplicaciones y adquisición*. Panamerican Health Organization. 1999.
- [Rodríguez, 2012] Rodríguez, A., et al. *Evaluación económica de la atención a pacientes en la epidemia de dengue*. MEDISAN [online]. 2012, vol.16, n.5. ISSN 1029-3019. [Consultado: mayo de 2012].
- [Suárez, 2000] Suárez L., et al. *Enfermedades emergentes y reemergentes: factores causales y vigilancia*. Rev Cubana Med Gen Integr. [online]. 2000, vol.16, n.6. ISSN 0864-2125. [Consultado: febrero de 2012].
- [Vega, 2008] Vega, L., et al. *La Inteligencia de Negocio. Su implementación mediante la Plataforma Pentaho*. Memorias I Taller de Proyección y Prospección en Tecnologías de Información y Comunicación Cuba – Venezuela. La Habana. 2008.



BIBLIOGRAFÍA

- Alba, D., et al. *Pasos para crear cubos con Mondrian Schema Workbench*. Open Bussiness Intelligence. 2009.
- Alonso, O. *BI: Inteligencia aplicada al negocio*. DAA Contenidos Digitales. S.L. 2005.
- Andrés, A. et al. *Vigila: un instrumento para la vigilancia epidemiológica de enfermedades transmisibles*. Rev Cubana Med Trop [online]. 1999, vol.51, n.1. ISSN 0375-0760. [Consultado: febrero de 2012].
- Bello, R. et al. *Modelos computacionales avanzados*. Departamento de Ciencias de la Computación. Universidad Central “Martha Abreu” de las Villas. Santa Clara, 1999.
- Bello R. et al. *Sistemas basados en el conocimiento: Teoría y Práctica*. M. E. En proceso de edición en la UCLV. Santa Clara. 2009.
- Bernabeu, R. D. *DATA WAREHOUSING: Investigación y Sistematización de Conceptos – HEFESTO: Metodología propia para la Construcción de un Data Warehouse*. Córdoba, Argentina. 2009.
- Boehm. *SW Cost Estimation with COCOMO II. Model Definition Manual*. Prentice Hall. 2000.
- Buchanan, L. et al. *Breve historia de la toma de decisiones*. Harvard Deusto Business Review. No.148. 2006.
- Cohen, D. *Sistemas de información para la toma de decisiones*. Segunda Edición. McGRAW – HILL. New York. 1996.
- Concepción, R. *La gestión de proyectos Informáticos sostenibles*. Tesis en opción al título de máster en Gestión Ambiental. Universidad de Holguín. 2004.
- Cruz, M. *El método Delphi en las investigaciones educacionales*. Editorial Academia. La Habana. 2009.



- Date, C. J. *Introducción a los sistemas de bases de datos*. Editorial Félix Varela, La Habana. 2003.
- Dirección de Informática del MINSAP. *Proyectos de investigación. Estadísticas*. Documento en línea. [Consultado: diciembre de 2009]. Disponible en: <http://www.di.sld.cu/url.php?url=dHV0b3JpYWxlcY5waHA=>.
- Estruch, L., et al. *Guías para la asistencia integral a pacientes con dengue*. Editorial Ciencias Médicas. La Habana. 2012.
- Harkey J., et al. *The essential client/server survival guide*. John Wiley & Sons. 1996.
- Hernández, J. M. *Software libre: técnicamente viable, económicamente sostenible y socialmente justo*. Infonomia, Red de Innovadores. 2005.
- Imhoff, C., et al. *Mastering data warehouse design. Relational and dimensional techniques*. Wiley Computer Publishing. Inc. Indianapolis. 2003.
- Inmon, W. H. *Building the data warehouse*. Third Edition. Wiley Computer Publishing. New York, 2002.
- Jacobson, I. et al. *El Proceso Unificado de Desarrollo de Software*. Addison – Wesley. 2000.
- Kouri, G. *El dengue, un problema creciente de salud en las Américas*. Rev Cubana Salud Pública [online]. 2011, vol.37, suppl.5. ISSN 0864-3466. [Consultado: febrero de 2012].
- Kimball, R. *Fact tables and dimension tables*. Documento en línea. [Consultado: mayo de 2010]. Disponible en: http://www.intelligententerprise.com/030101/602warehouse1_1.jhtml.
- Kimball, R., et al. *The Data Warehouse Toolkit. Second Edition. The Complete Guide to Dimensional Modeling*. Wiley Computer Publishing. New York. 2002.



- Larman, C. *Applying UML and Patterns. An Introduction to Object – Oriented Analysis and Design and the unified Process*. Second Edition. Prentice Hall. México. 2000.
- Marchesi, M., et al. *Extreme programming perspectives*. Addison Wesley. 2002.
- Mateu, C. *Desarrollo de aplicaciones Web*. Fundació per a la Universitat Oberta de Catalunya. Barcelona. 2004.
- McClure, C. R. *The information rich employee and information for decision making: Review and comments*. Information Processing and Management, 1978, vol. 14.
- MINSAP. *Estadísticas de Salud*. Editorial Pueblo y Educación. La Habana. 1985.
- MINSAP. *La Salud Pública en Cuba. Hechos y Cifras*. Dirección Nacional de Estadística. La Habana. 1999.
- MINSAP. *PNO Dengue*. Laboratorio Provincial SUMA. Centro Provincial de Higiene, Epidemiología y Microbiología. Santiago de Cuba. 2004.
- MINSAP. *Sistema de Información Estadística. Complementario de Salud. SIEC 2009*. Documento en línea. [Consultado: enero de 2009]. Disponible en: http://www.sld.cu/galerias/pdf/sitios/dne/siec_2009_completo.pdf.
- MINSAP. *Sistema de Vigilancia en Salud. Unidad de Análisis y Tendencias en Salud*. La Habana. 1995.
- Moss, L. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. 2003.
- Nader, J. *Sistema de Apoyo Gerencial Universitario*. Tesis de magister en Ingeniería del Software. Instituto Tecnológico de Buenos Aires. 2006.
- Pearson, W. *MDX at First Glance: Introduction to SQL Server MDX Essentials*. Documento en línea. [Consultado: mayo de 2010] Disponible en: <http://www.databasejournal.com/features/mssql/index.php>.



- Pence, N., Creeth, R. *An Introduction to OLAP*. Documento en línea. [Consultado: diciembre de 2008]. Disponible en: <http://www.OLAPreport.com/>.
- Pentaho. *Architecture. Layers of a Mondrian System*. Documento en línea. [Consultado: julio de 2009]. Disponible en: <http://mondrian.pentaho.org/documentation/architecture.php>.
- Pentaho. Open Source Business Intelligence. *Mondrian Schema Workbench*. Documento en línea. [Consultado: julio de 2009]. Disponible en: <http://wiki.pentaho.com/display/PMOPEN/Mondrian+Schema+Workbench>.
- Pentaho. Open Source Business Intelligence. *Creación de Soluciones Pentaho*. Pentaho Corporation. 2008.
- Pentaho. Open Source Business Intelligence. *Pentaho BI Suite Enterprise Edition*. Pentaho Corporation. 2008.
- Pentaho. Open Source Business Intelligence. *Pentaho Data Integration. Spoon 3.0 User Guide*. Pentaho Corporation. 2009.
- Pentaho. Open Source Business Intelligence. *Pentaho Report Design Wizard*. Pentaho Corporation. 2008.
- Poole, J., et al. *Common Warehouse Metamodel*. Wiley Computer Publishing. Inc. Indianapolis. 2003.
- Pressman, R. *Ingeniería del Software. Un enfoque práctico*. Quinta Edición. Mc Graw Hill. 2001.
- Purecos, L. *Sistemas inteligentes de soporte a la toma de decisiones*. Instituto Tecnológico de Monterrey. Documento en línea. [Consultado: septiembre de 2008]. Disponible en: <http://www.monografias.com/trabajos16/toma-de-decisiones/toma-de-decisiones.htm>.
- Rodríguez, A., et al. *Evaluación económica de la atención a pacientes en la epidemia de dengue*. MEDISAN [online]. 2012, vol.16, n.5. ISSN 1029-3019. [Consultado: mayo de 2012].



- Rodríguez, R., et al. *El establecimiento de sistemas de información en servicios de atención de salud: Guía para el análisis de requisitos, especificación de las aplicaciones y adquisición*. Panamerican Health Organization. 1999.
- Rojas, E. *DataMarts (Data Warehouse) Tool: Mondrian + JRubik*. ICIS Workshop. Centro Internacional de la Papa. 2006.
- Schmuller, J. *Aprendiendo UML en 24 horas*. Prentice Hall. México. 2000.
- Sparxs System. *Tutorial de UML*. Documento en línea. [Consultado: diciembre de 2008]. Disponible en: <http://www.sparxsystems.cl/bin/UML2SuperStructure.pdf>.
- Suárez L., et al. *Enfermedades emergentes y reemergentes: factores causales y vigilancia*. Rev Cubana Med Gen Integr. [online]. 2000, vol.16, n.6. ISSN 0864-2125. [Consultado: febrero de 2012].
- Thomsen, E. *OLAP Solutions. Building Multidimensional Information Systems*. Second Edition. Wiley Computer Publishing. New York. 2000.
- Vega, L., et al. *La Inteligencia de Negocio. Su implementación mediante la Plataforma Pentaho*. Memorias I Taller de Proyección y Prospección en Tecnologías de Información y Comunicación Cuba – Venezuela. La Habana. 2008.



ANEXOS

ANEXO 1. REGLAS DEL NEGOCIO

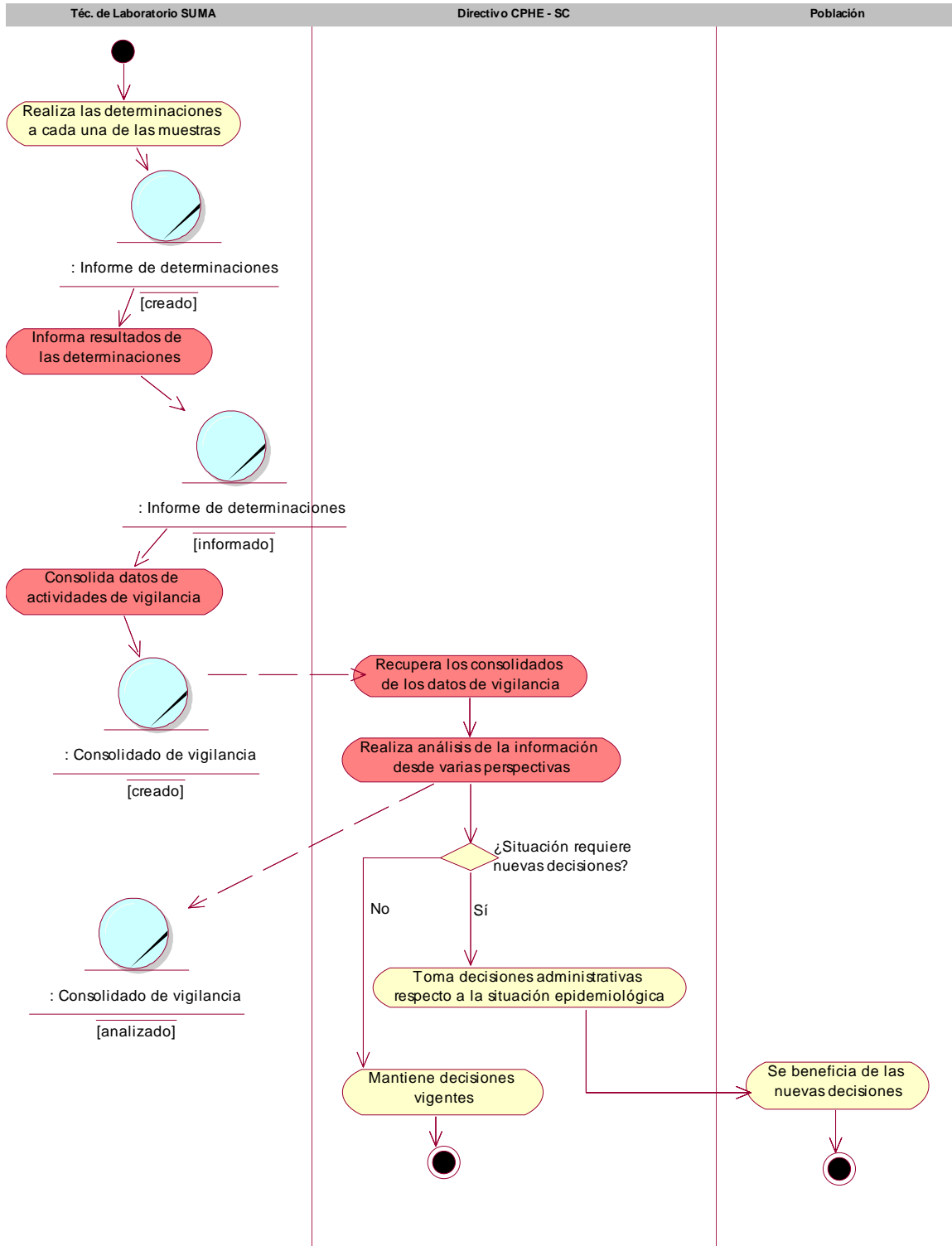
- Las muestras recibidas en el Laboratorio SUMA del CPHE son registradas y controladas, con la fecha en que fueron tomadas, así como los datos de la unidad en que fue tomada la misma, y los datos del paciente al que pertenece.
- El Sistema Nacional de Salud tiene tres niveles: nivel primario (áreas de salud), nivel secundario (hospitales generales y clínico – quirúrgicos) y nivel terciario (hospitales, centros e instituciones de salud especializados).
- La principal determinación a realizar es la prueba UMELISA Dengue IgM Plus (prueba de detección de anticuerpos IgM a los cuatro serotipos del virus del Dengue en suero mediante la tecnología SUMA).
- Las fechas siempre serán válidas.
- Para registrar una muestra debe haber sido registrado el paciente previamente.
- Para registrar los datos de una determinación, debe estar registrada previamente la muestra a la que se realiza.
- El paciente al que se le tomó la muestra y aún no se le ha realizado ninguna determinación, constituye un caso presuntivo, registrado o realizado.
- A cada muestra recibida se le debe realizar como mínimo dos determinaciones.
- El valor del corte (CUTT) o límite de decisión es un valor numérico brindado por la computadora o determinado de forma manual, que establece que aquellas muestras cuyo valor sea igual o superior a éste resulten positivas.
- El Border Line (BL) es un valor comprendido entre el valor de corte y un 15% por debajo de éste y que establece la denominada "zona gris".



- Una muestra se considera positiva si ambos duplicados (determinaciones) de la muestra arrojan resultados iguales o superiores al nivel de corte, o si el valor de fluorescencia de uno de los duplicados de la muestra cae en la zona gris y el otro por encima del nivel de corte.
- Una muestra se considera negativa si ambos duplicados (determinaciones) de la muestra caen por debajo de la zona gris, o si el valor de fluorescencia de uno de los duplicados cae en la zona gris y el otro cae por debajo de esta.
- Una muestra se considera BL si el valor de fluorescencia de ambos duplicados de la muestra cae en la zona gris.
- Una muestra reactiva es aquella que de forma repetida resulte positiva o BL al UMELISA Dengue IgM Plus.
- Toda muestra reactiva, constituye un caso confirmado.
- Se deberá repetir la prueba, si la fluorescencia de uno de los duplicados de la muestra cae por encima del nivel de corte y el otro cae por debajo de la zona gris.



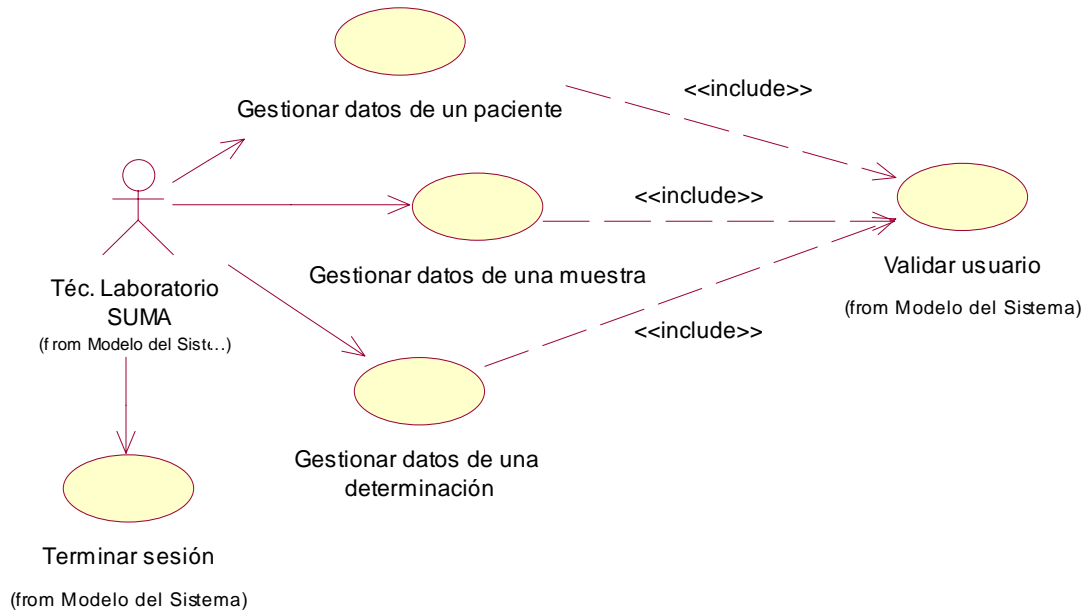
ANEXO 2. DIAGRAMA DE ACTIVIDADES PARA LA VIGILANCIA HIGIÉNICO – EPIDEMIOLÓGICA



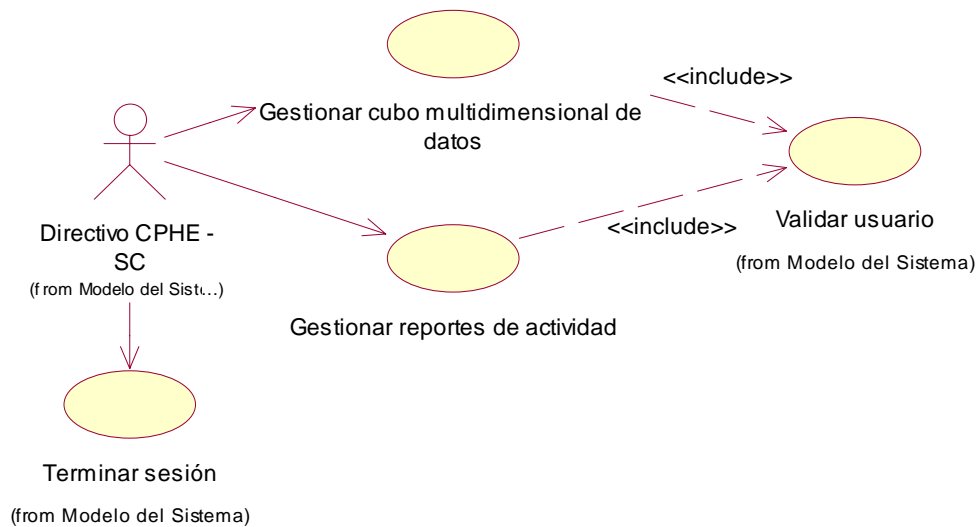


ANEXO 3. DIAGRAMAS DE CASOS DE USO DEL SISTEMA

Módulo de Procesamiento Transaccional en Línea

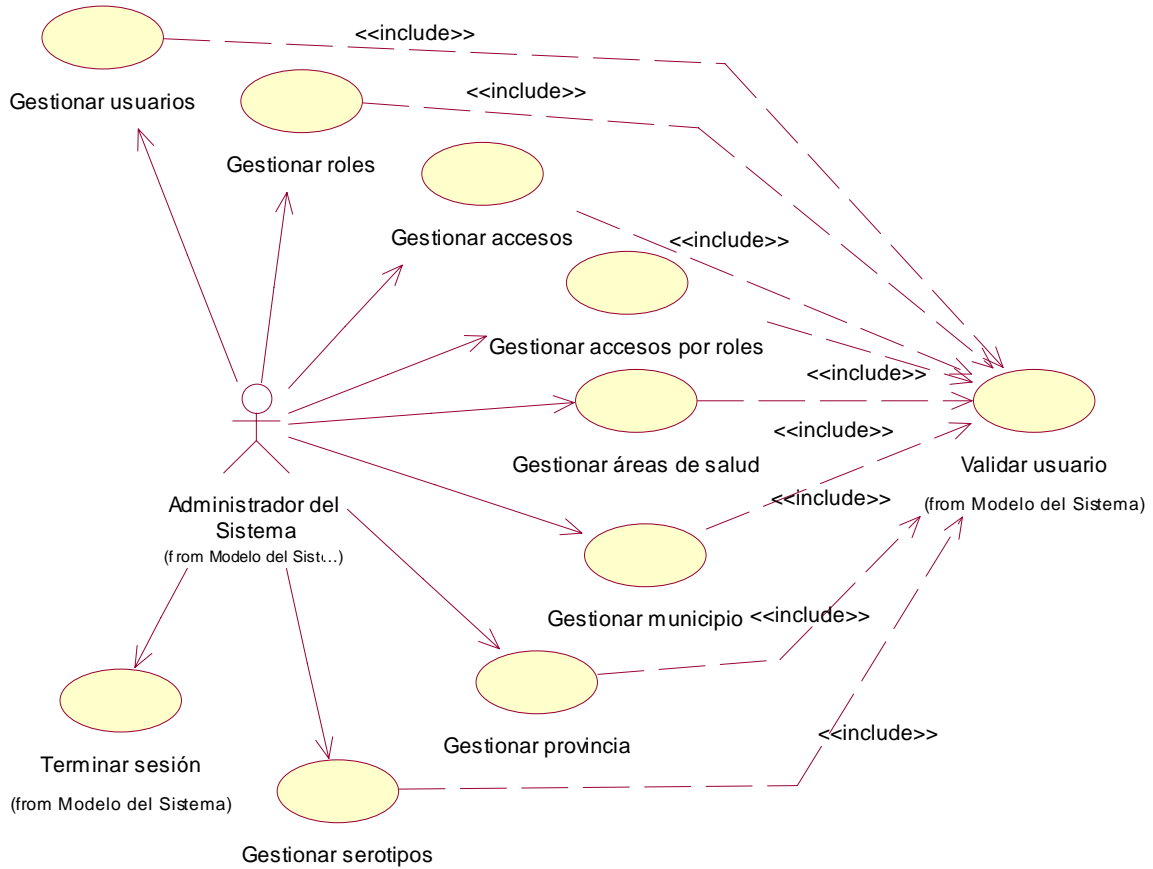


Módulo de Procesamiento Analítico en Línea





Módulo de Administración y Seguridad



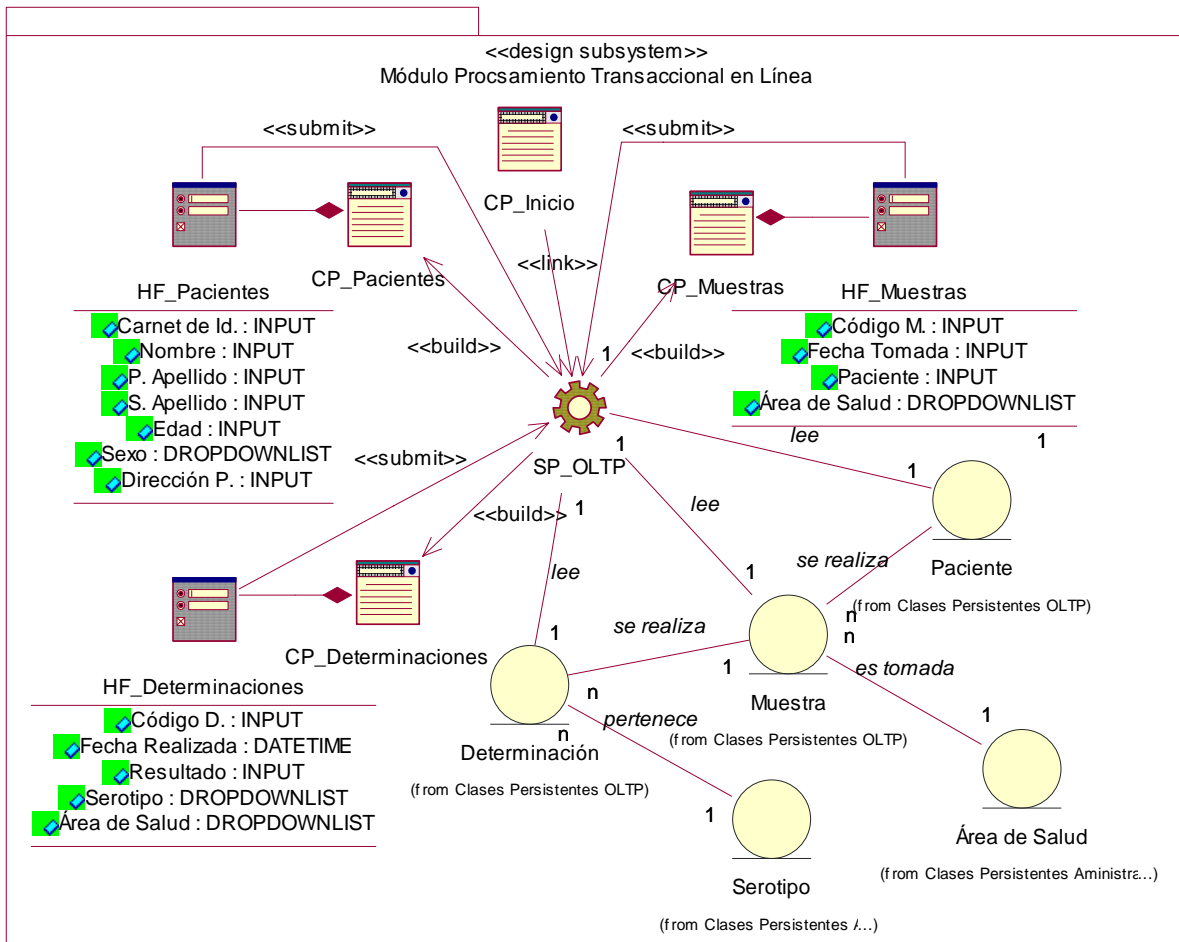
Módulo de Integración de los Datos





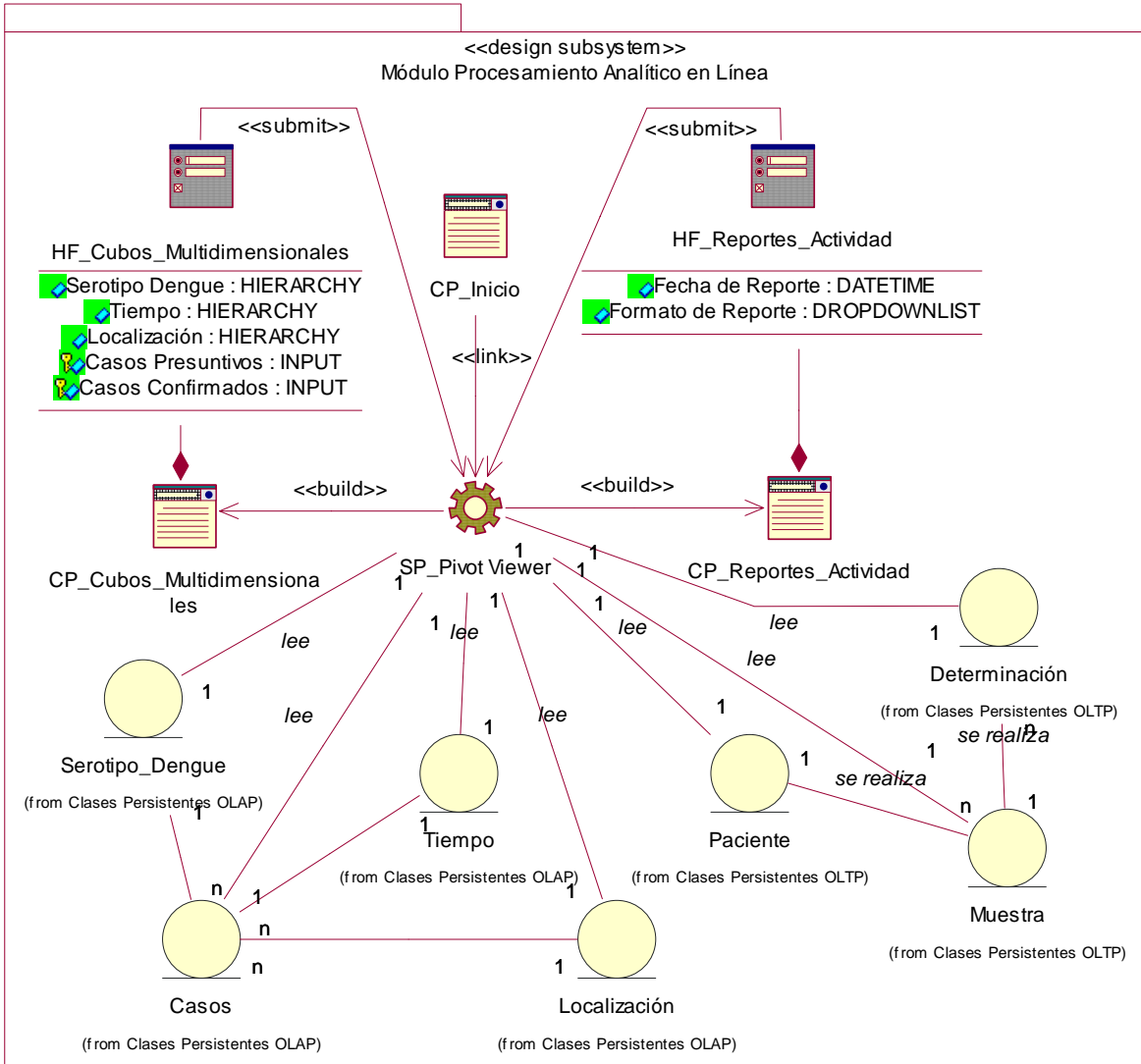
ANEXO 4. DIAGRAMAS DE CLASES

Módulo de Procesamiento Transaccional en Línea



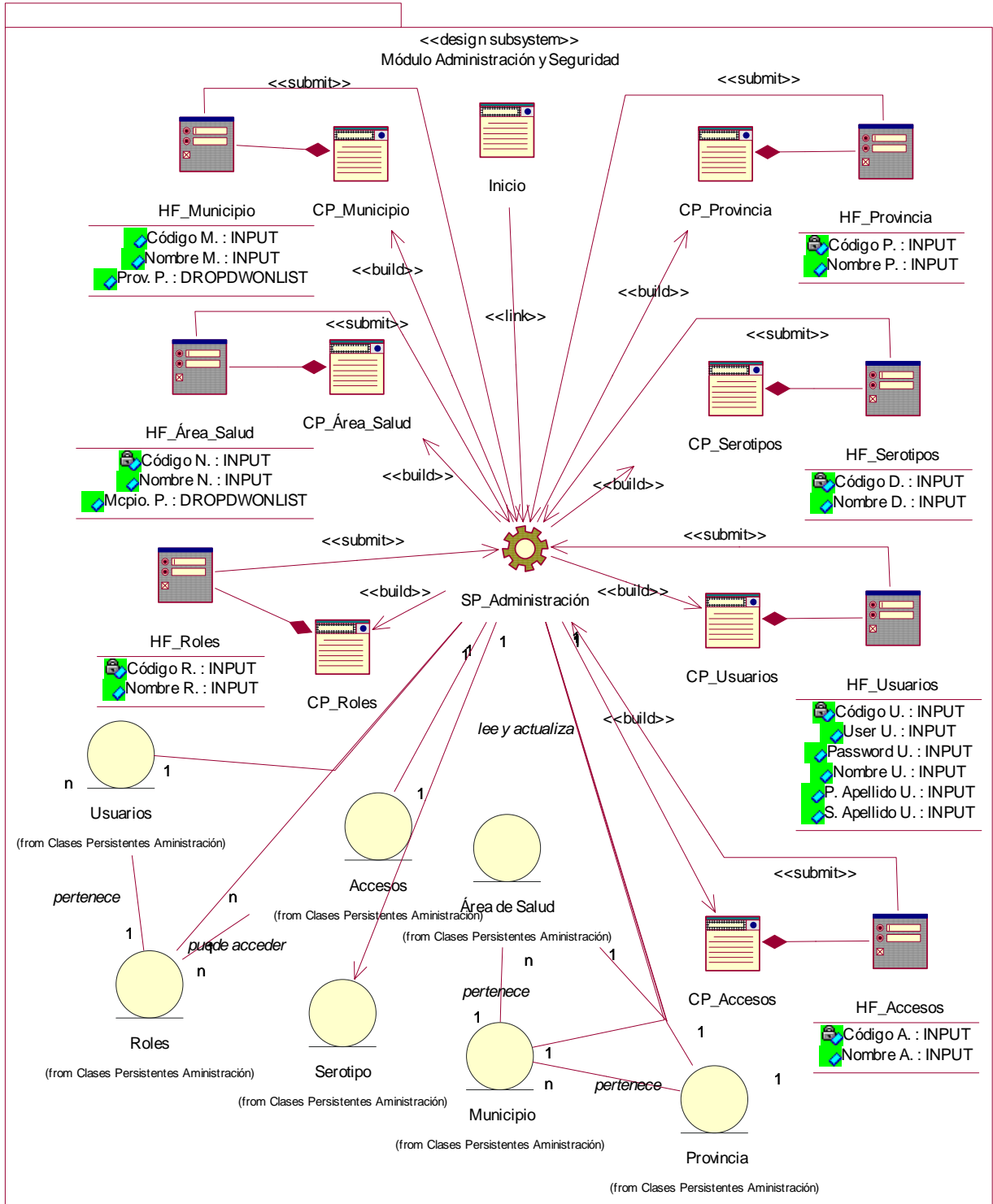


Módulo de Procesamiento Analítico en Línea





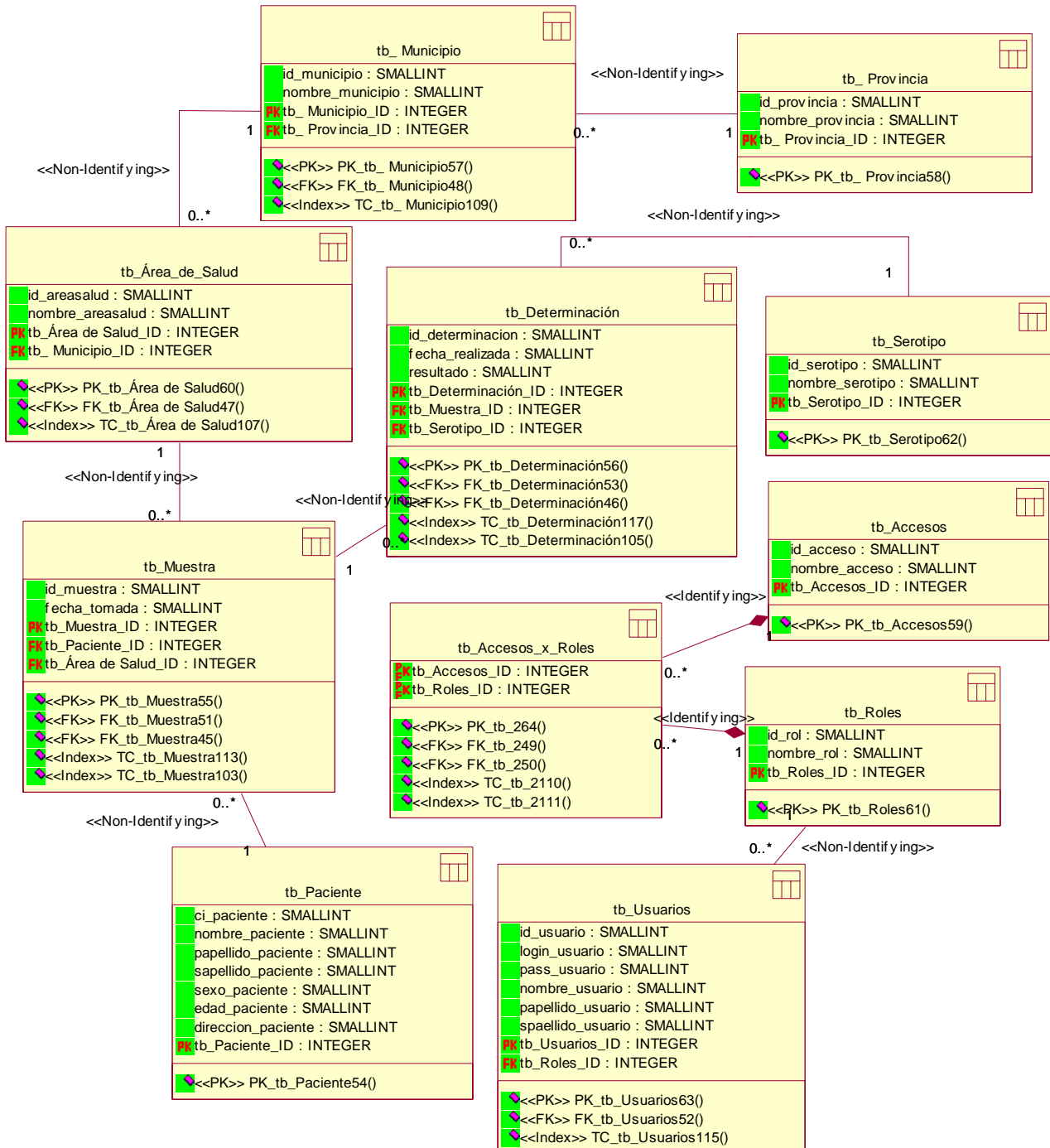
Módulo de Administración y Seguridad





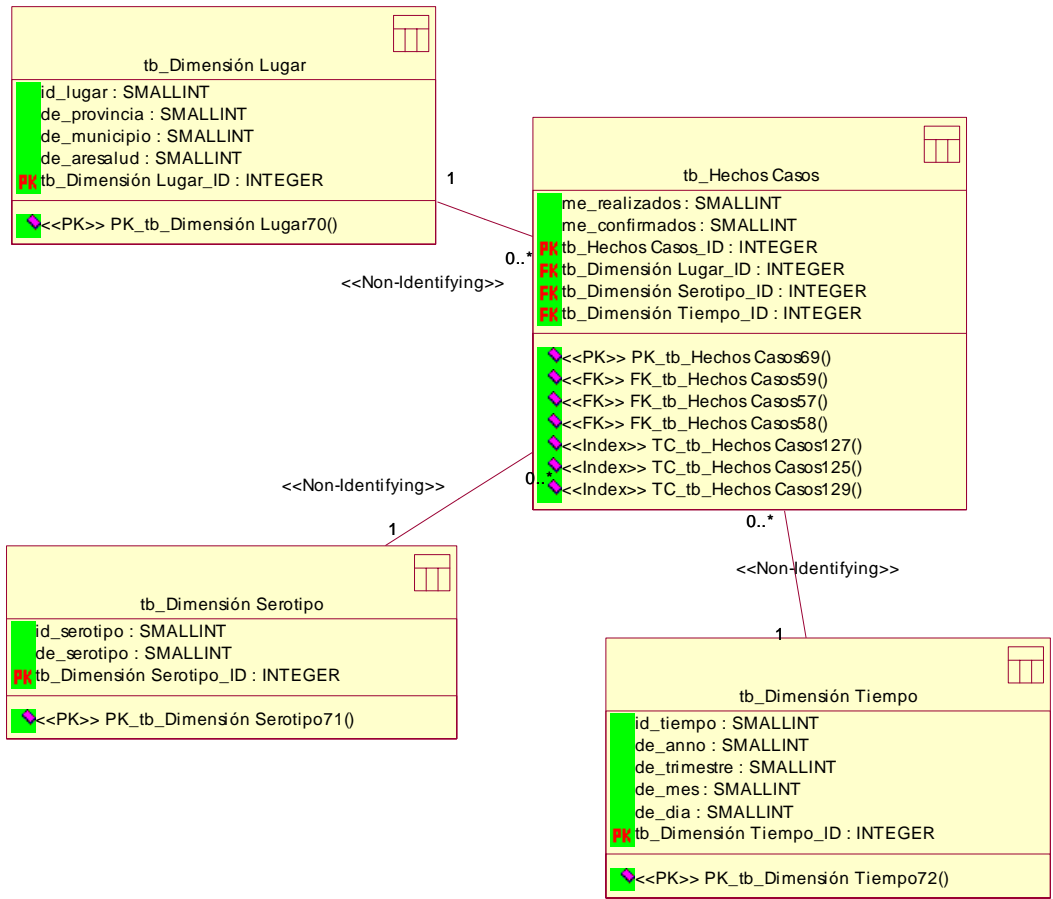
ANEXO 5. MODELOS DE DATOS

Base de Datos Operacional (bd_dengue_oltp)





Base de Datos Multidimensional (bd_dengue_olap)





ANEXO 6. PANTALLAS TÍPICAS DEL SISTEMA

epiDEN
Sistema de Procesamiento Transaccional
y Analítico para la Vigilancia
Epidemiológica del Dengue

Iniciar Sesión

Login

Contraseña

Usted ha sido autenticado en el Módulo de Procesamiento Transaccional en Línea [\[Salir \]](#)

Módulo Transaccional en Línea

- Pacientes
- Muestras
- Determinaciones

Gestionando Pacientes: 1 - 7 de 8

Filtro personalizado Todos los campos Sólo palabras completas

	C.I. Paciente	Nombre	P. Apellido	S. Apellido	Edad	Sexo	Dirección
	74122302031	Kenia	Charón	Díaz	38	F	Celda, Reparto Flores, Santiago de Cuba.
	74112302013	William	Infante	Pineda	38	M	Trinidad entre Sto Tomás y S. Félix
	60010125877	Yamilé	Ávila	Seco	52	F	Edificio 3, El Modelo, Caney, Santiago de Cuba
	60020256987	Aída	Aguilera	Vera	52	F	Altos del 30 de Noviembre, Reparto Militar, Santiago
	62030321475	Vivian	Walter	Sánchez	50	F	B - 9 Apto, 7, Dtto José Martí, Santiago
	70010125487	Vivian	Benito	Valenciano	42	F	Calle 6ta S/N Santa María, Santiago de Cuba
	52010125487	Cecilia	Joa	Ramos	60	F	A - 7 Apto, 1 Dtto Josçe Martí, Santiago de Cuba

1 2



Usted ha sido autenticado en el Módulo de Procesamiento Analítico en Línea [Salir]

Análisis de la Información

- Casos Realizados
- Casos Confirmados
- Casos Realizados y Confirmados

Reportes de Actividad

- Pacientes Registrados (Diario)
- Muestras Recibidas (Diario)
- Determinaciones Realizadas (Diario)

Lugar	Fecha	Medidas	
		● Casos Realizados	● Casos Confirmados
<input type="checkbox"/> Lugares	<input type="checkbox"/> Fechas	58	30
<input type="checkbox"/> Santiago de Cuba	<input type="checkbox"/> Fechas	58	30
	<input type="checkbox"/> 2012	58	30
<input type="checkbox"/> Santiago de Cuba	<input type="checkbox"/> Fechas	58	30
28 de Septiembre	<input type="checkbox"/> Fechas	1	0
	<input type="checkbox"/> 2012	1	0
30 de Noviembre	<input type="checkbox"/> Fechas	24	14
	<input type="checkbox"/> 2012	24	14
Camilo Torres	<input type="checkbox"/> Fechas	10	5
	<input type="checkbox"/> 2012	10	5
	<input type="checkbox"/> 1er Trimestre	10	5
Julán Grimau	<input type="checkbox"/> Fechas	22	10
	<input type="checkbox"/> 2012	22	10
	<input type="checkbox"/> 1er Trimestre	12	9
	<input type="checkbox"/> 2do Trimestre	10	1
Ramón López Peña	<input type="checkbox"/> Fechas	1	1

Usted ha sido autenticado en el Módulo de Procesamiento Analítico en Línea [Salir]

Análisis de la Información

- Casos Realizados
- Casos Confirmados
- Casos Realizados y Confirmados

Reportes de Actividad

- Pacientes Registrados (Diario)
- Muestras Recibidas (Diario)
- Determinaciones Realizadas (Diario)

Report Parameters

Save a Copy Print Email Search

LABORATORIO SUMA. CPHE - SANTIAGO DE CUBA

Reporte Diario de Determinaciones Realizadas

Determinaciones realizadas el día: 19.04.2012
 Fecha de emisión de este reporte: 10.05.2012

Cód. Muestra	Cód. Determinación	Serotipo	Fecha Realizada	Resultado
1	6	Serotipo 6	19 / 04 / 2012	455,00
1	7	Serotipo 6	19 / 04 / 2012	11,00
2	5	Serotipo 5	19 / 04 / 2012	13,96

Total de Determinaciones: 3

8,5 x 11 in 1 of 1



ANEXO 7. ESQUEMA MONDRIAN (XML) DONDE SE DEFINEN LOS CUBOS CREADOS

```
<Schema name="epiden" measuresCaption="">

  <Cube name="Casos Realizados" cache="true" enabled="true">
    <Table name="tb_hec_casos" schema="" alias="">
    </Table>
    <Dimension type="StandardDimension" foreignKey="fk_id_lugar" name="Lugar">
      <Hierarchy name="Lugar" hasAll="true" allMemberName="Lugares" primaryKey="id_lugar">
        <Table name="tb_dim_lugar" schema="" alias="">
        </Table>
        <Level name="Provincia" column="de_provincia" uniqueMembers="false">
        </Level>
        <Level name="Municipio" column="de_municipio" uniqueMembers="false">
        </Level>
        <Level name="rea de Salud" column="de_areasalud" uniqueMembers="false">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" foreignKey="fk_id_tiempo" name="Fecha">
      <Hierarchy name="Fecha" hasAll="true" allMemberName="Fechas" primaryKey="id_tiempo">
        <Table name="tb_dim_tiempo" schema="" alias="">
        </Table>
        <Level name="Año" column="de_anno" uniqueMembers="false">
        </Level>
        <Level name="Trimestre" column="de_trimestre" uniqueMembers="false">
        </Level>
        <Level name="Mes" column="de_mes" uniqueMembers="false">
        </Level>
        <Level name="Día" table="" column="de_dia" uniqueMembers="false">
        </Level>
      </Hierarchy>
    </Dimension>
    <Measure name="Casos Realizados" column="me_realizados" aggregator="sum" visible="true">
    </Measure>
  </Cube>

  <Cube name="Casos Confirmados" cache="true" enabled="true">
    <Table name="tb_hec_casos" schema="" alias="">
    </Table>
    <Dimension type="StandardDimension" foreignKey="fk_id_serotipo" name="Serotipo">
      <Hierarchy name="Serotipo" hasAll="true" allMemberName="Serotipos"
        primaryKey="id_serotipo">
        <Table name="tb_dim_serotipo" schema="" alias="">
        </Table>
        <Level name="Serotipo" column="de_serotipo" uniqueMembers="false">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" foreignKey="fk_id_lugar" name="Lugar">
      <Hierarchy name="Lugar" hasAll="true" allMemberName="Lugares" primaryKey="id_lugar">
        <Table name="tb_dim_lugar" schema="" alias="">
        </Table>
        <Level name="Provincia" column="de_provincia" uniqueMembers="false">
        </Level>
        <Level name="Municipio" column="de_municipio" uniqueMembers="false">
        </Level>
      </Hierarchy>
    </Dimension>
  </Cube>
```



```
</Level>
<Level name="&#193;rea de Salud" column="de_areasalud" uniqueMembers="false">
</Level>
</Hierarchy>
</Dimension>
<Dimension type="StandardDimension" foreignKey="fk_id_tiempo" name="Fecha">
<Hierarchy name="Fecha" hasAll="true" allMemberName="Fechas" primaryKey="id_tiempo">
<Table name="tb_dim_tiempo" schema="" alias="">
</Table>
<Level name="A&#241;o" column="de_anno" uniqueMembers="false">
</Level>
<Level name="Trimestre" column="de_trimestre" uniqueMembers="false">
</Level>
<Level name="Mes" column="de_mes" uniqueMembers="false">
</Level>
<Level name="D&#237;a" column="de_dia" uniqueMembers="false">
</Level>
</Hierarchy>
</Dimension>
<Measure name="Casos Confirmados" column="me_confirmados" aggregator="sum" visible="true">
</Measure>
</Cube>

<Cube name="Casos Realizados y Confirmados" cache="true" enabled="true">
<Table name="tb_hec_casos" schema="" alias="">
</Table>
<Dimension type="StandardDimension" foreignKey="fk_id_lugar" name="Lugar">
<Hierarchy name="Lugar" hasAll="true" allMemberName="Lugares" primaryKey="id_lugar">
<Table name="tb_dim_lugar" schema="" alias="">
</Table>
<Level name="Provincia" column="de_provincia" uniqueMembers="false">
</Level>
<Level name="Municipio" column="de_municipio" uniqueMembers="false">
</Level>
<Level name="&#193;rea de Salud" column="de_areasalud" uniqueMembers="false">
</Level>
</Hierarchy>
</Dimension>
<Dimension type="StandardDimension" foreignKey="fk_id_tiempo" name="Fecha">
<Hierarchy name="Fecha" hasAll="true" allMemberName="Fechas" primaryKey="id_tiempo">
<Table name="tb_dim_tiempo" schema="" alias="">
</Table>
<Level name="A&#241;o" column="de_anno" uniqueMembers="false">
</Level>
<Level name="Trimestre" column="de_trimestre" uniqueMembers="false">
</Level>
<Level name="Mes" table="" column="de_mes" uniqueMembers="false">
</Level>
<Level name="D&#237;a" column="de_dia" uniqueMembers="false">
</Level>
</Hierarchy>
</Dimension>
<Measure name="Casos Realizados" column="me_realizados" aggregator="sum" visible="true">
</Measure>
<Measure name="Casos Confirmados" column="me_confirmados" aggregator="sum" visible="true">
</Measure>
</Cube>

</Schema>
```



ANEXO 8. ENCUESTA PARA DETERMINAR EL COEFICIENTE DE COMPETENCIA DE LOS EXPERTOS

Nombre y apellidos: _____

Usted ha sido seleccionado como posible experto para ser consultado respecto al grado de relevancia de un sistema de procesamiento transaccional y analítico, que presumiblemente debe servir para ayudar a la toma de decisiones en el proceso de vigilancia epidemiológica del dengue.

Necesitamos, antes de realizarle la consulta correspondiente como parte del método “consulta a expertos”, determinar su coeficiente de competencia en este tema, a los efectos de reforzar la validez del resultado de la consulta. Por ello le rogamos que responda las siguientes preguntas de la forma más objetiva que le sea posible.

1. Marque con una cruz (X) en la tabla siguiente el valor que se corresponde con el grado de conocimientos que usted posee sobre el tema “Sistema de apoyo a la toma de decisiones”. Considere que la escala que le presentamos es ascendente, es decir, el conocimiento sobre el tema referido va creciendo desde 1 hasta 10.

1	2	3	4	5	6	7	8	9	10

2. Realice una autovaloración del grado de influencia que cada una de las fuentes que le presentamos a continuación ha tenido en su conocimiento y criterio sobre el tema “Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue”. Para ello marque con una cruz (X) según corresponda en A (alto), M (medio) o B (bajo).

Fuentes de argumentación	Grado de influencia de cada fuente		
	A	M	B
Análisis teórico realizado por usted.			
Su experiencia obtenida.			
Trabajo de autores nacionales.			
Trabajo de autores extranjeros.			
Su propio conocimiento del estado del problema en el extranjero.			
Su intuición.			

Gracias.



ANEXO 9. ENCUESTA APLICADA A LOS EXPERTOS

Nombre y apellidos: _____

Institución a la que pertenece: _____

Cargo actual: _____

Calificación profesional, grado científico o académico:

Profesor: _____ Licenciado: _____ Especialista: _____

Máster: _____ Doctor: _____

Años de experiencia en el cargo: _____

Dada la importancia que tiene la vigilancia epidemiológica del dengue, se está realizando una investigación que tiene como objetivo desarrollar un sistema informático orientado a técnicos, especialistas y administrativos de Higiene y Epidemiología, para el procesamiento de la información generada por esta actividad. El sistema está compuesto por tres grandes módulos: uno de procesamiento transaccional en línea de la información, uno para el procesamiento analítico en línea de la información, y otro para la administración y configuración del sistema y la seguridad del mismo. Precisamos de su evaluación y valoración, teniendo en cuenta diversos aspectos que a continuación le presentamos.

Indicaciones

A continuación le presentamos una tabla que contiene los aspectos de los módulos del ***Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue***, relacionados con el apoyo a la toma de decisiones por parte de los administrativos. A la derecha aparece la escala:

MR: Muy relevante. BR: Bastante relevante. R: Relevante.

PR: Poco relevante. NR: No relevante.

Marque con una cruz (X) en la celda que se corresponda con el grado de relevancia que usted otorga a cada aspecto del sistema de referencia.

Le agradecemos anticipadamente el esfuerzo que sabemos hará para responder la presente encuesta, con la mayor fidelidad posible a su manera de pensar.



Sobre el Sistema de Procesamiento Transaccional y Analítico para la Vigilancia Epidemiológica del Dengue					
Aspectos	Relevancia				
	MR	BR	R	PR	NR
1. Gestión de datos de pacientes.					
2. Gestión de datos de muestras.					
3. Gestión de datos de determinaciones.					
4. Gestión de la seguridad del sistema.					
5. Implementación de funcionalidades de búsqueda y ordenación.					
6. Gestión de datos de elementos del sistema.					
7. Flexibilidad y mantenimiento del sistema.					
8. Integración de los datos entre las bases de datos del sistema.					
9. Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del Dengue.					
10. Facilidades para la manipulación de los cubos multidimensionales.					
11. Reorientar los cubos multidimensionales de datos.					
12. Aumentar nivel de agregación en los datos de los cubos multidimensionales.					
13. Disminuir nivel de agregación en los datos los cubos multidimensionales.					
14. Generación de reportes parametrizados.					
15. Flexibilidad y prestaciones de la plataforma utilizada.					

Escriba a continuación qué aspectos deben ser incluidos o eliminados, en caso de que así lo considere:

Aspectos que se proponen ser incluidos

Aspectos que se proponen ser eliminados

Escriba a continuación qué aspectos deben cambiados en caso de que así lo considere:

Donde dice

Debe decir

Muchas gracias.



ANEXO 10. RESULTADOS GENERALES DE LA ENCUESTA APLICADA A LOS EXPERTOS

Tabla – Frecuencia Absoluta

Aspectos	MR	BR	R	PR	NR	Total
1. Gestión de datos de pacientes.	1	0	24	0	0	25
2. Gestión de datos de muestras.	1	2	22	0	0	25
3. Gestión de datos de determinaciones.	1	3	21	0	0	25
4. Gestión de la seguridad del sistema.	8	15	2	0	0	25
5. Implementación de funcionalidades de búsqueda y ordenación.	9	12	4	0	0	25
6. Gestión de datos de elementos del sistema.	4	18	3	0	0	25
7. Flexibilidad y mantenimiento del sistema.	19	4	2	0	0	25
8. Integración de los datos entre las bases de datos del sistema.	17	7	1	0	0	25
9. Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del Dengue.	17	7	1	0	0	25
10. Facilidades para la manipulación de los cubos multidimensionales.	20	2	3	0	0	25
11. Reorientar los cubos multidimensionales de datos.	8	13	4	0	0	25
12. Aumentar nivel de agregación en los datos de los cubos multidimensionales.	8	11	6	0	0	25
13. Disminuir nivel de agregación en los datos los cubos multidimensionales.	7	12	6	0	0	25
14. Generación de reportes parametrizados.	21	1	3	0	0	25
15. Flexibilidad y prestaciones de la plataforma utilizada.	6	15	4	0	0	25

Tabla – Inverso de la Frecuencia Absoluta Acumulada

Aspectos	MR	BR	R	PR
1. Gestión de datos de pacientes.	0,04	0,04	1,00	1,00
2. Gestión de datos de muestras.	0,04	0,12	1,00	1,00
3. Gestión de datos de determinaciones.	0,04	0,16	1,00	1,00
4. Gestión de la seguridad del sistema.	0,32	0,92	1,00	1,00
5. Implementación de funcionalidades de búsqueda y ordenación.	0,36	0,84	1,00	1,00



6. Gestión de datos de elementos del sistema.	0,16	0,88	1,00	1,00
7. Flexibilidad y mantenimiento del sistema.	0,76	0,92	1,00	1,00
8. Integración de los datos entre las bases de datos del sistema.	0,68	0,96	1,00	1,00
9. Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del Dengue.	0,68	0,96	1,00	1,00
10. Facilidades para la manipulación de los cubos multidimensionales.	0,8	0,88	1,00	1,00
11. Reorientar los cubos multidimensionales de datos.	0,32	0,84	1,00	1,00
12. Aumentar nivel de agregación en los datos de los cubos multidimensionales.	0,32	0,76	1,00	1,00
13. Disminuir nivel de agregación en los datos los cubos multidimensionales.	0,28	0,76	1,00	1,00
14. Generación de reportes parametrizados.	0,84	0,88	1,00	1,00
15. Flexibilidad y prestaciones de la plataforma utilizada.	0,24	0,84	1,00	1,00

Tabla – Determinación de los Puntos de Corte

Aspectos	MR	BR	R	PR	Prom	Valores de escala
1. Gestión de datos de pacientes.	-1,75	-1,75	3,49	3,49	0,87	0,95
2. Gestión de datos de muestras.	-1,75	-1,17	3,49	3,49	1,01	0,81
3. Gestión de datos de determinaciones.	-1,75	-0,99	3,49	3,49	1,06	0,76
4. Gestión de la seguridad del sistema.	-0,47	1,41	3,49	3,49	1,98	-0,16
5. Implementación de funcionalidades de búsqueda y ordenación.	-0,36	0,99	3,49	3,49	1,90	-0,08
6. Gestión de datos de elementos del sistema.	-0,99	1,17	3,49	3,49	1,79	0,03
7. Flexibilidad y mantenimiento del sistema.	0,71	1,41	3,49	3,49	2,27	-0,45
8. Integración de los datos entre las bases de datos del sistema.	0,47	1,75	3,49	3,49	2,30	-0,48
9. Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del Dengue.	0,47	1,75	3,49	3,49	2,30	-0,48
10. Facilidades para la manipulación de los cubos multidimensionales.	0,84	1,17	3,49	3,49	2,25	-0,43
11. Reorientar los cubos multidimensionales de datos.	-0,47	0,99	3,49	3,49	1,88	-0,06
12. Aumentar nivel de agregación en los	-0,47	0,71	3,49	3,49	1,80	0,02



datos de los cubos multidimensionales.						
13. Disminuir nivel de agregación en los datos los cubos multidimensionales.	-0,58	0,71	3,49	3,49	1,78	0,04
14. Generación de reportes parametrizados.	0,99	1,17	3,49	3,49	2,29	-0,47
15. Flexibilidad y prestaciones de la plataforma utilizada.	-0,71	0,99	3,49	3,49	1,82	0,00
Puntos de corte	-0,39	0,69	3,49	3,49	1,82	

Tabla – Conclusiones Generales

Aspectos	MR	BR	R	PR	NR
1. Gestión de datos de pacientes.	-	-	Sí	-	-
2. Gestión de datos de muestras.	-	-	Sí	-	-
3. Gestión de datos de determinaciones.	-	-	Sí	-	-
4. Gestión de la seguridad del sistema.	-	Sí	-	-	-
5. Implementación de funcionalidades de búsqueda y ordenación.	-	Sí	-	-	-
6. Gestión de datos de elementos del sistema.	-	Sí	-	-	-
7. Flexibilidad y mantenimiento del sistema.	Sí	-	-	-	-
8. Integración de los datos entre las bases de datos del sistema.	Sí	-	-	-	-
9. Inclusión de técnicas de análisis OLAP para la vigilancia epidemiológica del Dengue.	Sí	-	-	-	-
10. Facilidades para la manipulación de los cubos multidimensionales.	Sí	-	-	-	-
11. Reorientar los cubos multidimensionales de datos.	-	Sí	-	-	-
12. Aumentar nivel de agregación en los datos de los cubos multidimensionales.	-	Sí	-	-	-
13. Disminuir nivel de agregación en los datos los cubos multidimensionales.	-	Sí	-	-	-
14. Generación de reportes parametrizados.	Sí	-	-	-	-
15. Flexibilidad y prestaciones de la plataforma utilizada.	-	Sí	-	-	-

