



**UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS  
FACULTAD 2**

**TRABAJO DE DIPLOMA PARA OPTAR POR EL TÍTULO DE INGENIERO EN  
CIENCIAS INFORMÁTICAS**

**COMPONENTE PARA EL ANÁLISIS DE PROCESOS APLICANDO LA  
PERSPECTIVA TEMPORAL EN EL SISTEMA DE INFORMACIÓN  
HOSPITALARIA DEL CENTRO DE INFORMÁTICA MÉDICA**

**AUTORES**

Lisandra Castañeda Domínguez

Antonio Valladares Martínez

**TUTOR**

Ing. Arturo Orellana García

**La Habana, 24 de junio de 2015**

“Año 57 de la Revolución”



*“Los estudiantes son en su mayoría revolucionarios. Revolucionarios por naturaleza, porque pertenecen a ese estrato de jóvenes que se abren a la vida y que adquieren todos los días conocimientos nuevos”*

*Ernesto “Che” Guevara*

## DECLARACIÓN DE AUTORÍA

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los \_\_\_\_ días del mes de \_\_\_\_\_ del año \_\_\_\_\_.

**Lisandra Castañeda Domínguez**

---

Firma del Autor

**Antonio Valladares Martínez**

---

Firma del Autor

**Arturo Orellana García**

---

Firma del Tutor

## DATOS DE CONTACTO

Ing. Arturo Orellana García ([aorellana@uci.cu](mailto:aorellana@uci.cu)): graduado de Ingeniería en Ciencias Informáticas en la Universidad de las Ciencias Informáticas en el año 2012. Labora en el Centro de Informática Médica (CESIM). Se desempeña como tutor de las asignaturas PP4 a PP7 y ha sido analista de los proyectos SIAPS, Telemedicina Domiciliaria y Ensayos Clínicos del Centro de Informática Médica. Ha tutorado 2 trabajos de diploma. Actualmente es estudiante de maestría.

## AGRADECIMIENTOS

### **Lisandra Castañeda Domínguez**

*Le agradezco a mi mamá (la persona más tierna y amorosa que existe) por estar siempre conmigo ayudándome y apoyándome, por confiar siempre en mí y por exigirme ser cada vez mejor, por escucharme y por ser mi mamita.*

*A mi papá, por todo su amor y cariño, por ser el papa más lindo que existe, por estar siempre a mi lado, por el he tratado de ser una persona de la cual pueda estar orgullosa siempre.*

*A mi hermanito, porque a pesar de tener siempre opiniones diferentes, es uno de los seres que más amo en esta vida.*

*A mi esposo, por todo su amor, su comprensión, por estar siempre a mi lado, por reír conmigo en los momentos felices y secar mis lágrimas en los días tristes, por ser mí amigo, por ser mi amorcito.*

*A mi abue, por ser como yo digo, mi Calviño, por ser mi confidente, por aconsejarme y por darme aliento para seguir adelante en los momentos más difíciles de mi carrera, por sacrificarse tanto por mí y dar lo mejor de ella para que yo sea feliz.*

*A toda mi familia, a mis tíos, mis tías, mis primos, a mi abuelita Margarita, por toda su preocupación, porque cada vez que he necesitado su ayuda siempre han estado ahí presentes para mí.*

*A mis suegros, que se son como unos padres para mí, por ayudarme y por preocuparse tanto de mí.*

*A Marcos que fue uno de las personas que más me ayudo en los años más difíciles de mi carrera, por aconsejarme tanto, por darme aliento y por confiar en mí.*

*A mi compañero de tesis, por soportarme todo este tiempo, por escogerme y por confiar en mí, para realizar este trabajo tan importante en nuestras vidas.*

*A todos mis amigos, a mis compañeros de aula y a todos los profesores que formaron parte de esta etapa tan bonita de mi vida.*

*A mi tutor, por soportarnos todos estos meses y ayudarnos a ser mejores, por toda su preocupación y ayuda.*

*A todos los del grupo de investigación de Minería de procesos por su apoyo y su ayuda.*

*A todos de corazón gracias.*

### **Antonio Valladares Martínez**

*A mi mamá y a mi papá por siempre tener confianza en mí, por todo lo que han hecho para que yo sea la persona que soy hoy.*

*A mi hermano y mis abuelas.*

*A mi novia por siempre estar ahí cuando más la necesitaba, por estar a mi lado en los momentos buenos pero sobre todo por aguantar todo este tiempo a mi lado lo cual fue un reto enorme por el estrés que yo emanaba, también por calmarme en los momentos de desespero.*

*A mi compañera de tesis por haber accedido a realizar la tesis conmigo y aguantar mis jaranas.*

*Al tutor por estar siempre en combate junto a nosotros.*

*A Yoandry González Castro.*

*A mis amigos de la vieja escuela (La gente del proyecto) y a los nuevos que he hecho en la universidad.*

*A todos los del grupo de investigación de Minería de procesos, al profesor Damián.*

*A mis profesores de estos 5 años*

## DEDICATORIA

### ***Lisandra Castañeda Domínguez***

Dedico el presente trabajo de diploma a mis padres, a mis abuelas, a mi esposo, a mi hermano y a mi familia por todo su amor y su apoyo.

### ***Antonio Valladares Martínez***

Dedico el trabajo de diploma a mi madre, a mi padre, mis abuelas, mi hermano y mi novia.

---

## RESUMEN

La Minería de procesos incluye diferentes perspectivas, entre las que se encuentra la perspectiva temporal, la cual resulta ser efectiva y útil en sectores críticos como la salud, pues permite determinar las diferencias de tiempo dentro de un proceso asistencial y detectar cuellos de botella, los cuales ralentizan el funcionamiento de los procesos en las instituciones hospitalarias. El objetivo de la presente investigación es desarrollar un componente para la obtención de modelos de procesos aplicando la perspectiva temporal de Minería de procesos, a partir del registro de eventos del Sistema de Información Hospitalaria del Centro de Informática Médica. Para el desarrollo del componente se utiliza como lenguaje de programación: Java en su versión 1.6, como Entorno Integrado de Desarrollo: Eclipse 3.4.2 y se utilizó PostgreSQL 9.3.3 como Sistema Gestor de Base de Datos. El resultado de la presente investigación es la personalización del plugin *Variants Miner* la cual muestra modelos de procesos en forma de árbol, donde las actividades que conforman el proceso, contienen su duración máxima, media, mínima y total. Los resultados de la aplicación del componente en el sistema para el cual fue desarrollado, permitirá obtener a los analistas de procesos criterios cualitativos y cuantitativos para la toma de decisiones clínico-administrativas y permitirá contar con un mecanismo para generar nuevos conocimientos, a partir de la interpretación de los datos almacenados en el sistema.

**Palabras clave:** Centro de Informática Médica, Minería de procesos, perspectiva temporal, Sistema de Información Hospitalaria.

**ABSTRACT**

*Mining process includes different perspectives, among which is the temporal perspective, which proves to be effective and useful in critical sectors such as health, it allows determining the time differences in a care process and identify bottlenecks. The bottlenecks slow down the operation of processes in hospitals. The objective of this research is to develop a component for producing models using temporal perspective of mining processes, from event log of the Hospital Information System of Medical Informatics Center. For the development of the component is used as the programming language Java in version 1.6, Integrated Development Environment: Eclipse 3.4.2 and PostgreSQL 9.3.3 as the Database Manager System. The result of this research is the personalization of the Variants Miner plugin which displays models of processes in a tree, where the activities involved in the process, contain its minimum, average, maximum and total duration. The results of the application of the component in the system for which it was developed, will allow to processes analysts to get qualitative and quantitative criteria for making clinical management decisions and will provide a mechanism to generate new knowledge from the interpretation of data stored in the system.*

**Keywords:** *Center for Medical Informatics, Hospital Information System, Process Mining, time perspective.*

## ÍNDICE GENERAL

Introducción.....	1
Capítulo 1: Fundamentación teórica de la investigación .....	6
1.1.    Conceptos fundamentales.....	6
1.1.1. <i>Procesos</i> .....	6
1.1.2. <i>Gestión basada en procesos</i> .....	6
1.1.3. <i>Análisis de procesos de negocio</i> .....	7
1.1.4. <i>Minería de procesos</i> .....	8
1.1.5. <i>Subproceso</i> .....	11
1.1.6. <i>Árbol de variantes</i> .....	12
1.1.7. <i>Registro de eventos</i> .....	12
1.1.8. <i>Soluciones existentes a nivel internacional y nacional</i> .....	14
1.2.    Tecnologías y herramientas .....	19
1.2.1. <i>Marcos de trabajo</i> .....	19
1.2.2. <i>Metodología de desarrollo</i> .....	20
1.2.3. <i>Lenguajes</i> .....	20
1.2.4. <i>Tecnologías a utilizar</i> .....	21
1.2.5. <i>Herramientas a utilizar</i> .....	23
Capítulo 2: Propuesta de solución.....	26
2.1.    Propuesta de solución.....	26
2.2.    Patrones arquitectónicos.....	27
2.3.    Patrones de diseño .....	28
2.3.1. <i>Patrón de diseño GRASP</i> .....	28
2.4.    Técnica de Minería de variantes.....	29
2.4.1. <i>Etapas de la Minería de variantes</i> .....	31
2.5.    Desarrollo del componente para la visualización de modelos de procesos aplicando la perspectiva temporal de Minería de procesos.....	33
2.5.1. <i>Componente aplicando la perspectiva temporal de Minería de procesos</i> .....	34
2.5.2. <i>Personalización del plugin Variants Miner</i> .....	35
2.5.3. <i>Integración de la personalización del plugin Variants Miner en el HIS del CESIM, para la visualización de modelos de procesos</i> .....	35
Capítulo 3: Validación de la solución propuesta .....	39

3.1. Comprensión del modelo generado por el componente desarrollado en el HIS.....	39
3.1.1. Identificación de métricas para el análisis de procesos aplicando la perspectiva temporal	43
3.2. Validación de las actividades .....	44
Conclusiones.....	48
Recomendaciones.....	49
Referencias Bibliográficas .....	50
Bibliografía .....	53
Anexos .....	57

## ÍNDICE DE FIGURAS

Figura 1.1: Representación de los tipos de técnicas de Minería de procesos.....	8
Figura 1.2: Vista de análisis de proceso utilizando DISCO. ....	15
Figura 1.3: Resultado de la técnica "Analyze using Dotted Chart". ....	17
Figura 2.1: Flujo de información de la propuesta de solución .....	26
Figura 2.2: Esquema representativo de la técnica Minería de variantes. ....	30
Figura 2.2: Variantes de proceso.....	30
Figura 2.3: Variantes contenidas en un árbol de variantes. ....	31
Figura 2.4: Fórmula de media aritmética. ....	34
Figura 2.5: Área para la entrada de datos. ....	36
Figura 2.6: Área de resultados. Interfaz del "modelo" generado por el componente ....	38
Figura 3.1: Modelo del proceso Procesar Solicitudes .....	40
Figura 3.2: Modelo del subproceso crear_cronograma_de_entrega.....	41
Figura 3.3: Modelo del subproceso crear_contrato.....	42
Figura 3.4: Plugin Add <i>Identities to log</i> de ProM.....	45
Figura 3.5: Cantidad de actividades del proceso Solicitar Productos.....	46
Figura 3.6: Modelo del proceso Solicitar Productos.....	46
Figura anexo 2: Diagrama de casos de uso Generar modelo de proceso. ....	57

### INTRODUCCIÓN

Las organizaciones actuales, independientemente de su tamaño y las actividades que realizan, se enfrentan a los mercados competitivos en los que pueden conciliar la satisfacción de los clientes, con la eficiencia económica<sup>1</sup> de sus actividades. Esto provoca la necesidad de desarrollar habilidades para el mejoramiento de los procesos y de esta forma elevar su calidad en la organización.

Una estrategia que permite mejorar y controlar de forma continua los procesos y las interrelaciones entre ellos, es el enfoque basado en procesos. Como soporte a esta forma organizativa se han desarrollado un conjunto de sistemas de información (Alfonso, 2014), lo que ha contribuido de manera eficiente a la gestión de los procesos y que la información que se genere esté más organizada.

Los sistemas de información orientados a procesos, cuya implantación resulta de vital importancia en las organizaciones, requieren de un modelado explícito de los procesos de negocio de las mismas, pero dicho modelado no es algo trivial, pues se necesita un amplio conocimiento de los mismos para modelarlos con la suficiente exactitud. En muchas ocasiones se carece de este conocimiento, por lo que su adquisición conlleva a un largo procedimiento, pues estos modelos de procesos suelen acabar representando más bien cómo debería llevarse a cabo el proceso, que cómo se está ejecutando en realidad. (Dios, Framiñán, Domínguez, & León, 2010)

Una alternativa en ascenso para modelar la ejecución real de los procesos es la Minería de procesos, disciplina de investigación relativamente joven que se ubica entre la inteligencia computacional y la minería de datos, por una parte, y la modelación y análisis de procesos, por otra. El objetivo de la Minería de procesos es descubrir, monitorear y mejorar los procesos reales a través de la extracción de conocimiento de los registros de eventos, ampliamente disponibles en los actuales sistemas de información. (Yzquierdo, 2013)

---

<sup>1</sup> Eficiencia: Según el Diccionario de la Real Academia Española, eficiencia (del latín *efficientia*) es la capacidad de disponer de alguien o de algo para conseguir un efecto determinado.

Eficiencia económica: como la eficiencia con la cual un sistema económico utiliza los recursos productivos a fin de satisfacer sus necesidades.

Para el análisis de procesos se siguen diferentes perspectivas entre las que se encuentran la organizacional, la de casos y la temporal, siendo esta última una de las más importantes debido a que se relaciona con la ocurrencia y frecuencia de los eventos. Cuando los eventos tienen asociados marcas de tiempo es posible descubrir cuellos de botella<sup>2</sup>, medir niveles de servicio, monitorear la utilización de recursos y predecir el tiempo de procesamiento restante de casos en ejecución (Manifiesto sobre Minería de Proceso, 2011).

El análisis de procesos desde la perspectiva temporal permite, tener una mejor idea de los procesos con los que se están trabajando, los autores del artículo Modelado y análisis de un proceso quirúrgico mediante técnicas de Minería de procesos, desarrollaron un estudio acerca del análisis temporal de procesos (Dios, Framiñán, Domínguez, & León, 2010), permitió descubrir a partir de técnicas de Minería de procesos que, la duración media de algunas actividades eran relativamente largas, logrando realizar representaciones de modelos teniendo en cuenta las características temporales de cada una de las actividades.

En investigaciones realizadas por los autores del trabajo Procedimiento de Conformidad aplicando Minería de procesos desde la perspectiva tiempo en el Sistema de Información Hospitalaria del Centro de Informática Médica (Cruz & Perdomo, 2014), detectaron retrasos en el cumplimiento de diferentes instancias de algunas actividades que se llevan a cabo (i.e. Proceso Solicitar Productos y Emergencias), desviaciones en el patrón de comportamiento y actividades que representan cuellos de botella en el funcionamiento del sistema.

En sectores críticos como el de la salud, el análisis de procesos desde la perspectiva temporal resulta ser especialmente efectivo y útil (van der Aalst et. al, 2011), pues determina el tiempo promedio de atención a un paciente, las diferencias de tiempo dentro de un proceso asistencial para un grupo de pacientes con el mismo diagnóstico y las actividades dentro del hospital que puedan representar cuellos de botella, ya sea por tardanza en la atención a los pacientes o por sobrecarga en los servicios médicos.

Con el objetivo de automatizar los procesos del nivel secundario de salud en Cuba, en el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI), se desarrolló el Sistema de Información Hospitalaria (HIS, por sus siglas en inglés), orientado a

---

<sup>2</sup> Cuellos de botella: es un límite en la capacidad de transferencia de información de un sistema o una conexión, que puede reducir el tráfico en condiciones de sobrecarga.

satisfacer las necesidades de: almacenamiento, procesamiento, recopilación e interpretación de los datos clínico-administrativos que se generan en las instituciones hospitalarias.

El HIS, cuenta con un componente para la extracción y transformación de las trazas de ejecución en registros de eventos, el cual almacena información con marcas de tiempo de las actividades, como fecha de inicio, fecha de fin y duración de las mismas. Entre las operaciones que se pueden realizar con la información contenida en el Registro de eventos se encuentra:

- Calcular la duración media y el tiempo total de las actividades, así como su desviación típica.
- Obtener resultados de los tiempos de demora de ejecución de las actividades.
- Conocer qué productos o insumos tarda más o menos en ser despachados o enviados a las labores asistenciales.
- Obtener información sobre el desempeño de la organización o las actividades de proceso.

Además de recopilar y almacenar datos, el HIS debe procesar e interpretar los mismos, sin embargo, no cuenta con funcionalidades que permitan realizar análisis de tiempo (mencionados anteriormente) en la ejecución de las actividades de proceso en el sistema. La información que posee el registro de eventos del sistema no está siendo aprovechada, privando a los analistas de procesos e investigadores el acceso esta fuente de conocimientos sin explotación. Esta situación limita el uso de nuevos elementos para enriquecer criterios que favorecerían la toma de decisiones sobre el rendimiento de la institución y la ejecución de sus actividades.

Por lo anteriormente planteado se identifica como **problema a resolver**: ¿cómo obtener modelos de procesos aplicando la perspectiva temporal de Minería de procesos, a partir del registro de eventos del Sistema de Información Hospitalaria del Centro de Informática Médica?

El problema está enmarcado en el **objeto de estudio**: gestión basada en procesos en instituciones hospitalarias, centrado en el **campo de acción**: análisis de procesos aplicando la perspectiva temporal en instituciones hospitalarias.

Para solucionar el problema planteado, se define como **objetivo general**: desarrollar un componente para la obtención de modelos de procesos aplicando la perspectiva temporal de Minería de procesos al registro de eventos del Sistema de Información Hospitalaria del Centro

de Informática Médica, que permita obtener criterios cualitativos y cuantitativos para la toma de decisiones clínico-administrativas.

Se proponen las siguientes tareas de la investigación:

1. Analizar los sistemas que a nivel nacional e internacional, permiten realizar análisis de procesos aplicando la perspectiva temporal de Minería de procesos.
2. Analizar las herramientas a utilizar, así como de los principales algoritmos y técnicas de la Minería de procesos.
3. Desarrollar un componente aplicando la perspectiva temporal de Minería de procesos al registro de evento del Sistema de Información Hospitalaria del CESIM.
4. Personalizar el plugin *Variants Miner*, incorporándole el componente desarrollado y definiendo las configuraciones necesarias para el Sistema de Información Hospitalaria.
5. Desarrollar un componente para la visualización de modelos de procesos, obtenidos a partir de la personalización del plugin *Variants Miner*.
6. Validar la solución propuesta a partir de un caso de estudio.

Los **métodos científicos** utilizados para desarrollar la investigación (Hernández, 2011) fueron:

### Métodos teóricos:

- Análisis histórico-lógico: se utilizó para analizar el surgimiento, trayectoria y evolución de la Minería de procesos para una mejor comprensión del objeto y campo de estudio.
- Analítico -sintético: se puso en práctica realizando un estudio de las principales técnicas y herramientas existentes que son utilizadas en la aplicación de Minería de procesos.
- Inductivo-deductivo: se utilizó en la aplicación de casos de pruebas al sistema, llegando a conclusiones a partir de las respuestas proporcionadas por este.

### Métodos empíricos:

- Observación: como instrumento para adquirir conocimiento sobre el campo de acción a través de la investigación directa de las herramientas.

El documento está estructurado en tres capítulos, siendo estos:

El **Capítulo 1. Fundamentación teórica de la investigación**, trata los conceptos fundamentales sobre los diferentes elementos en la que se basa la presente investigación, incluye un análisis del estado del arte, a nivel nacional e internacional de sistemas que permitan realizar análisis de procesos desde una perspectiva temporal y se hace una

descripción de las tecnologías y herramientas usadas para el desarrollo del componente propuesto.

El **Capítulo 2. Propuesta de solución**, describe la propuesta de solución, junto a la modelación del flujo de información del componente ha desarrollado, se describen los patrones arquitectónicos y los de diseños utilizados; se realiza una explicación de la técnica utilizada y del procedimiento seguido para el desarrollo del componente.

Por último, en el **Capítulo 3. Validación de la solución propuesta**, se valida la aplicación, a partir de un caso de estudio.

## **CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA DE LA INVESTIGACIÓN**

En el presente capítulo se abordan los conceptos fundamentales sobre la gestión basada en procesos, una reseña acerca de la Minería de procesos, sus principales técnicas, algoritmos y perspectivas. Se analizan las principales herramientas de Minería de procesos que permite realizar análisis de procesos aplicando la perspectiva temporal y se describen las tecnologías y herramientas utilizadas para dar cumplimiento al objetivo general de la presente investigación.

### **1.1. Conceptos fundamentales**

#### **1.1.1. Procesos**

La palabra proceso aparece del latín *processus*, que significa avance y progreso.

Según la serie de normas internacionales ISO 9000 (2005) se define un proceso como “conjunto de actividades mutuamente relacionadas o que interactúan, las cuales transforman elementos de entrada en resultados”.

El modelo europeo de excelencia empresarial, conocido como modelo EFQM, define proceso como, una secuencia de actividades que va añadiendo valor mientras se produce un determinado producto o servicio a partir de determinadas aportaciones.

Autores como Ihaddadene (Ihaddadene, 2008), Aytulun (Aytulun, 2008) definen que un proceso de negocio es una descripción o conjunto de actividades de una organización en términos de tareas, agentes, reglas y procedimientos.

Mientras que Pappa (Pappa, Faltin, & Zimmermann, 2009) plantea que un proceso de negocio se presenta como un marco para las actividades donde participan e interactúan diversos actores para producir un producto o un servicio y así lograr los objetivos y metas definidos por una empresa. Los autores de la presente investigación se basan para la realización de la misma, en la definición de proceso dada por la serie de normas internacionales ISO 9000 (2005).

#### **1.1.2. Gestión basada en procesos**

La serie de normas internacionales ISO 9001 (2008) promueve la adopción de un enfoque basado en procesos cuando se desarrolla, implementa y mejora la eficacia de un sistema de

gestión de la calidad, para aumentar la satisfacción del cliente mediante el cumplimiento de sus requisitos.

Algunos autores plantean que la Gestión de Procesos de Negocio, (*Business Process Management*, BPM por sus siglas en inglés) apoya la definición, ejecución y seguimiento de los procesos de negocio por medio de sistemas que soportan el análisis, diseño, modelado, automatización y promulgación de estos mismos, permitiendo el cumplimiento de reglas de negocio y políticas, así como la asignación de recursos. (Wojtkowski, 2009)

Por su parte Weske (Weske, van der Aalst, & H.M.W, 2004) plantea para BPM como un sistema que busca: “Apoyar los procesos de negocio utilizando métodos, técnicas y software para diseñar, aprobar, controlar y analizar los procesos operativos con interacción de seres humanos, organizaciones, aplicaciones, documentos y otras fuentes de información”.

### 1.1.3. Análisis de procesos de negocio

El análisis de procesos de negocio (*Business Process Analysis*, BPA por sus siglas en inglés) tiene como objetivo investigar aquellas propiedades de los procesos de negocios que no son evidentes, ni triviales. El término análisis en BPA se utiliza de una manera muy amplia e incluye la simulación, diagnóstico, verificación y análisis de desempeño o rendimiento de los procesos. (Vergidis, 2008)

En trabajos generales sobre BPA se pueden encontrar técnicas basadas en análisis estadístico simple, datos y técnicas de Minería de procesos, redes neuronales, heurísticas, algoritmos genéticos entre otros.

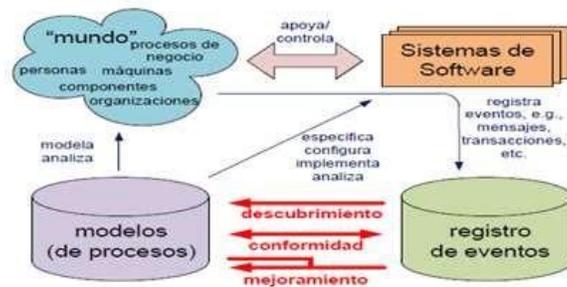
En el libro *Best Practice Measurement Strategies* (Hernández, 2012) se presenta un caso específico de organizaciones médicas, donde expresan que realizar BPA resulta ser difícil debido a la naturaleza de los procesos del sector salud, que se caracterizan por tener una alta dinámica, ser complejos, ad-hoc<sup>3</sup> y multidisciplinarios. De esta manera se propone realizar el estudio implementando técnicas de Minería de procesos, ya que éstas prometen una alta comprensión sobre el análisis de los datos registrados en el sistema de información.

---

<sup>3</sup> Ad-hoc significa "para este propósito" o "para esto". Es una frase latina que a menudo se utiliza para indicar que un determinado acontecimiento es temporal y es destinado a ese propósito específico.

## 1.1.4. Minería de procesos

La Minería de procesos enfoca sus esfuerzos en la extracción de información útil y relevante sobre los procesos de la organización, tomando como punto de partida los datos contenidos en los registros de eventos. (Pérez, Yzquierdo, Silverio, & Lazo, 2013) Existen tres tipos fundamentales de técnicas de minería de proceso, el descubrimiento de proceso, la verificación de conformidad y el mejoramiento de modelos, como se muestra en la **Fig. 1.1**.



**Figura 1.1:** Representación de los tipos de técnicas de Minería de procesos

Fuente: van der Aalst, 2011.

El descubrimiento de procesos abarca el conjunto de técnicas para la obtención de un modelo de procesos partiendo de un registro de eventos. Para muchas organizaciones resulta sorprendente la capacidad de los algoritmos de descubrimiento para identificar los procesos reales partiendo únicamente de los registros de eventos asociados a su ejecución. (van der Aalst, 2011)

En el chequeo o verificación de conformidad se compara un modelo de proceso existente con un registro de eventos del mismo proceso. La verificación de conformidad puede ser usada para chequear si la realidad, tal como está almacenada en el registro de eventos, es equivalente al modelo y viceversa. (IEEE Task Force On Process Mining, 2011)

La mejora de modelos de procesos es extender o mejorar un modelo de proceso existente usando la información acerca del proceso real almacenada en algún registro de eventos.

Los autores de la presente investigación se centrarán en el descubrimiento de procesos, pues permite determinar a partir de un registro de eventos, un modelo de proceso, que describe el comportamiento presente en el mismo.

## Descubrimiento de procesos

El descubrimiento es el área más atendida dentro de la minería de proceso, diseñándose varios algoritmos, con diferentes enfoques. Se entiende por algoritmo de descubrimiento una función que mapea un registro de eventos hacia un modelo de proceso [van der Aalst, W.M.P., 2011]. Existen varios algoritmos desarrollados tales como:

- **Alineación de trazas:** la alineación de trazas facilita el problema de determinar patrones interesantes en medianos y grandes registros, los que se determinan automáticamente y se muestran al usuario. Las trazas alineadas constituyen una representación de las tareas de acuerdo a un orden relativo y su estructuración en casos. El orden establecido entre las tareas permite identificar los patrones de control de flujo que se manifiestan en el proceso (Bose et al., 2012).
- **Minería heurística:** es una técnica que aplica algoritmos basados en heurística. Estos algoritmos descubren casi todas las estructuras comunes de control (decisión, secuencia, lazos, paralelismo) y son robustos ante ruidos en el log de eventos. (Weijters et al., 2006).
- **Minería difusa:** es una técnica que muestra las actividades y sus relaciones según diferentes niveles de abstracción. Utiliza similitudes con la forma de representar la información en la cartografía, aplicando conceptos como: abstracción, personalización, agregación y énfasis. Dicha técnica produce un modelo basado en grafos, donde existen dos tipos de nodos, los que representan una actividad y los que agrupan un conjunto de actividades denominados clústeres. (Günther and van der Aalst, W.M.P., 2007)
- **Minería de variantes:** es una técnica que permite construir diferentes variantes de modelos de proceso, a partir de la descomposición en subprocesos. La descomposición en subprocesos se realiza utilizando patrones de control de flujo donde se obtienen diferentes variantes de descomposición, de cada subproceso, descartando el comportamiento poco frecuente y/o asumiendo la presencia de comportamiento, en un ámbito definido por parámetros de configuración. (Alfonso, 2014)

## Justificación de la selección de la técnica Minería de variantes

Los autores de la presente investigación seleccionan la técnica de descubrimiento Minería de variantes, con el objetivo de obtener modelos procesos, pues esta técnica permite la identificación de los patrones de control de flujo, así como la obtención de una vista global y jerárquica del proceso, permite además la descomposición en subprocesos a partir de la combinación de los comportamientos presentes y ausentes en el registro de eventos, por lo que resulta una de las técnicas más completas del descubrimiento de proceso.

### **Perspectivas para aplicar Minería de procesos**

La Minería de procesos puede abarcar diferentes perspectivas (Leoni, W.M.P, & van der Aalst, 2012) para realizar el análisis de los datos. Estas perspectivas describen el tipo de información que se puede extraer dependiendo de la información almacenada en el registro de eventos o de los objetivos que se persiguen con la aplicación de las técnicas de Minería de procesos. A continuación se describen algunas de las perspectivas más utilizadas:

- **Perspectiva organizacional:** responde a la pregunta ¿Quién realiza un trabajo determinado? trabajando sobre la información referente a los usuarios del sistema que ejecutaron cada actividad en específico. El objetivo es estructurar la organización al clasificar a las personas en términos de las funciones y roles, así como mostrar la relación entre los distintos usuarios (construir una red social) (Medeiros, 2006).
- **Perspectiva de control de flujo:** se enfoca en el control de flujo, i.e., el orden de ejecución de las actividades. El objetivo es encontrar una buena caracterización de todos los caminos posibles. El resultado se expresa típicamente en términos de una red de Petri<sup>4</sup> o alguna otra notación de procesos (i.e. EPCs, BPMN, o diagramas de actividad UML), actores (i.e. personas, sistemas, o departamentos) están involucrados y cómo se relacionan. El objetivo es ya sea estructurar la organización clasificando las personas en términos de roles y unidades organizacionales, o mostrar la red social. (IEEE Task Force On Process Mining, 2011)
- **Perspectiva de casos:** se enfoca en las propiedades de los casos. Se deben analizar las propiedades de los casos que sean relevantes para el estudio. Es útil cuando existen

---

<sup>4</sup> Una red de Petri es un grafo orientado con dos tipos de nodos: lugares (representados mediante circunferencias) y transiciones (representadas por segmentos rectos verticales).

desviaciones y se requiere estudiar en qué punto del proceso se originó. Por ejemplo, si se detectan rutas que no corresponden a lo establecido, se pueden analizar los casos en que ello ocurre (Chamorro & Valderr, 2013).

- **Perspectiva temporal:** puede ser aplicada cuando se tiene información con relación al tiempo en el registro de eventos, permitiendo integrar al modelo las marcas de tiempo, la misma se relaciona con la ocurrencia y frecuencia de los eventos. Cuando los eventos tienen asociados marcas de tiempo, es posible descubrir cuellos de botella, medir niveles de servicio, monitorear la utilización de recursos, y predecir el tiempo de procesamiento restante de casos en ejecución. (van der Aalst, 2011)

### Justificación de la selección de la perspectiva temporal

Se selecciona la perspectiva temporal, pues permite analizar todo aquello referente al tiempo que pueda ser útil para representar de una forma más precisa el proceso real, como por ejemplo, la duración media de las tareas, su desviación típica, la suma total de su duración. Además permite obtener una representación del modelo teniendo en cuenta las características de cada una de las actividades en lo que a tiempo se refiere y de esta forma poder hacerse una mejor idea de los procesos con los que se está trabajando.

### 1.1.5. Subproceso

Un subproceso es una encapsulación de las actividades del negocio la cual representa una unidad de trabajo lógica y coherente. Los subprocesos poseen sus propios atributos y metas, pero contribuyen a alcanzar la meta general del proceso. Un subproceso es también un proceso y la mínima expresión de un subproceso es una actividad.

Un proceso se puede descomponer en múltiples subprocesos utilizando los siguientes patrones de control de flujo:

- **Secuencia ( $w_s$ ):** dos subprocesos se ejecutan secuencialmente si uno ocurre inmediatamente después del otro.
- **Selección exclusiva ( $w_x$ ):** dos subprocesos forman parte de una selección exclusiva si, en un punto de decisión, se puede ejecutar solamente uno de ellos.
- **Selección no Exclusiva ( $w_o$ ):** dos subprocesos son opciones de una selección no exclusiva si, en un punto de decisión, pueden ejecutarse ambos o solamente uno de

ellos.

- **Paralelismo ( $w_p$ ):** dos subprocesos se ejecutan en paralelo si ambos se ejecutan simultáneamente.
- **Lazo ( $w_l$ ):** dos subprocesos se encuentran en un lazo si se repiten múltiples veces.

Cada repetición comienza con la ejecución del primer subproceso (**Do**), continúa con el segundo (**Redo**) y termina con el **Do**. El **Redo** puede ser un subproceso vacío, por lo que el único repetido sería el **Do**.

Se denota como  $S_P$  el conjunto de los subprocesos del proceso  $P$  y como  $W = \{w_s, w_x, w_o, w_p, w_l\}$  el conjunto de los patrones de control de flujo. Un  $s_i \in S_P$  se puede descomponer a través de diferentes  $w | \forall w_j \in W$  en varios  $s_k | \forall k \neq i : s_k \in S_P$ , hasta el nivel de actividad. Se construye entonces un árbol en el que el mayor nivel de abstracción se encuentra en la raíz y el menor en las hojas (Alfonso, 2014).

### 1.1.6. Árbol de variantes

Un árbol de variantes es una representación de diferentes descomposiciones en subprocesos aplicadas al proceso  $P$  a partir de un registro de eventos. Este árbol está compuesto por dos tipos de nodos, los nodos subproceso y los nodos patrón. Un nodo subproceso representa un subproceso y posee tantos nodos patrones de hijos, como posibles descomposiciones se hayan identificado para el subproceso. Un nodo patrón representa una descomposición de su padre, de acuerdo a un patrón de control de flujo, por lo que un nodo patrón posee dos o más nodos subproceso como hijos.

El nodo raíz es un nodo subproceso y se refiere a todo el proceso. Los nodos hojas son siempre nodos de tipo subproceso. Las aristas del árbol representan caminos dentro del proceso. (Alfonso, 2014). La presente investigación se centrará en los nodos hojas del árbol de variantes y no en las aristas del mismo.

### 1.1.7. Registro de eventos

En el glosario de términos del Manifiesto sobre Minería de procesos (IEEE Task Force On Process Mining, 2011) se define evento como una acción almacenada en el registro, por

ejemplo, el inicio, conclusión o cancelación de una actividad para una instancia particular de un proceso.

Cada evento en un registro utiliza un identificador mientras cada atributo de los eventos puede ayudar a extender el modelo con información extra, tales como el recurso que ejecuta o inicia la actividad, la marca de tiempo del evento, o elementos de datos registrados con el evento. Dependiendo de la cantidad o tipo de información disponible será la perspectiva que se podrá asumir y la información que se podrá extraer.

El punto de partida para cualquier actividad de Minería de procesos son los eventos registrados. Nos referimos a colecciones de eventos como registros de eventos, sin embargo, esto no implica que los eventos deban estar almacenados en archivos de registros dedicados. Los eventos podrían estar almacenados en tablas de bases de datos, registros de mensajes, archivos de correo, registros de transacciones, y otras fuentes de datos. (IEEE Task Force On Process Mining, 2011).

El tipo de dato en el que generalmente son definidos los registros de eventos es el Xlog y uno de los estándares que los estructuran es el XES (*eXtensible Event Stream*). El principal propósito de este estándar es ofrecer un formato de intercambio de registros de eventos entre herramientas y dominios de aplicaciones (Günther, 2009).

Para asegurar un análisis de Minería de procesos exitoso, además del formato de almacenamiento del registro de eventos se debe garantizar su calidad. La misma se define a partir de tres aspectos fundamentales: confiabilidad, completitud y seguridad.

La confiabilidad consiste en que los eventos deben ser confiables, es decir, debería ser seguro asumir que los eventos registrados realmente ocurrieron y que los atributos de los eventos son correctos. La completitud se relaciona a que los registros de eventos deberían ser completos, dado un determinado contexto, no puede faltar ningún evento. Además, cualquier evento registrado debe tener una semántica bien definida. Por otra parte si los datos de eventos son seguros se tienen en cuenta consideraciones de privacidad y seguridad al registrar los eventos. (IEEE Task Force On Process Mining, 2011)

### 1.1.8. Soluciones existentes a nivel internacional y nacional

#### Nivel internacional

A nivel internacional existen una serie de herramientas que permiten realizar análisis de procesos aplicando la perspectiva temporal, las mismas son: *ProM*, *Disco*, *Dotted chart analysis*, *PPM Chart*, *ARIS Process Performance Manager*, entre otros.

A continuación se describen las herramientas antes mencionadas.

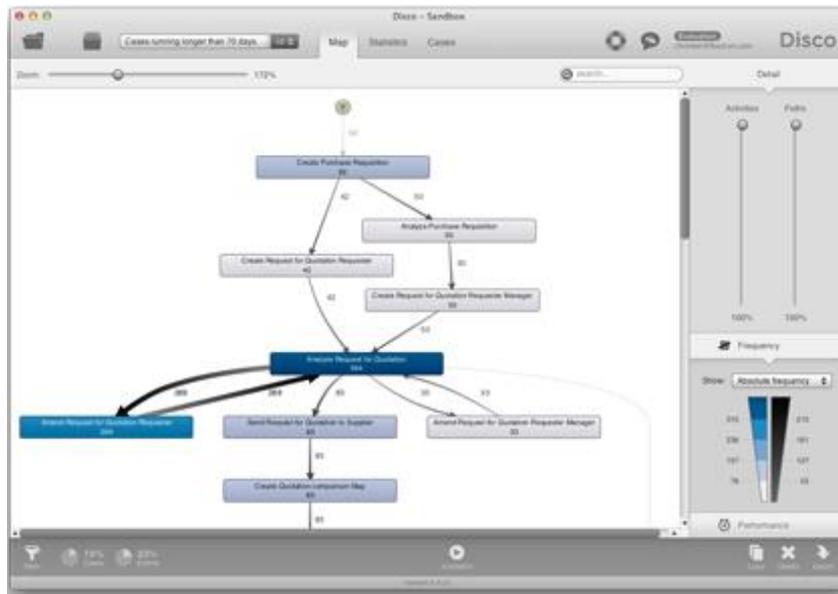
- **Herramientas líderes**

**Disco:** es una solución desarrollada y distribuida por *Fluxicon*, es una herramienta que tiene fines comerciales y posee licencia comercial y académica. Sin embargo, presenta una buena alternativa para personas que tienen conocimientos básicos de Minería de procesos, pero que no son necesariamente expertas, ya que concentra las mejores técnicas (se basa en *Fuzzy Mining*), las que ya han sido probadas y han demostrado buenos resultados, facilitando la labor del analista. Además, permite simular, usar distintos niveles de generalización y tiene una interfaz de usuario atractiva e intuitiva. (Van der Aalst, 2009)

Funcionalidades de DISCO (Rozinat, 2012):

1. Resumen estadísticas (Número de eventos, casos, actividades, recursos y atributos; inicio y marca de hora final del registro)
2. Resumen gráficos con superposición interactiva
  - Eventos en el tiempo.
  - Casos activos en el tiempo.
  - Variantes de casos.
  - Duración de los Casos (tiempo de producción).
  - La duración media de las actividades.
  - La media de los tiempos de espera.
3. Lista de los casos con número de eventos, marca de tiempo de inicio, fin y la duración.
4. Lista de variantes con número de casos, el número de eventos, duración media y el rango duración.
5. Actividad, recursos, que atribuyen estadísticas (número de valores; mínima, máxima y media de frecuencia; desviación estándar de frecuencia).

6. Actividad y de recursos gráficos con superposición interactiva.
  - Frecuencia.
  - La media de duración.
  - Gama duración.
  - Duración total.
7. Vista detallada de la tabla con la lista ordenada de actividades / recursos y su frecuencia absoluta y relativa, duración media y el rango de duración.



**Figura 1.2:** Vista de análisis de proceso utilizando DISCO.

Fuente: Rozinat, 2012.

**Marco de trabajo ProM:** es una herramienta genérica, de código abierto que agrupa las principales técnicas desarrolladas en la Minería de procesos. Su desarrollo es impulsado por la Universidad Tecnológica de Eindhoven, Holanda, donde se encuentra el grupo de investigación más destacado en el área. ProM tiene una arquitectura que permite ir agregando progresivamente nuevos *plugins*<sup>5</sup>. Además, soporta una amplia gama de modelos de flujo de control, incluyendo varios tipos de redes de Petri, *Event-driven Process Chains* (EPC), *Business Process Modeling Notation* (BPMN) y *Business Process Execution Language* (BPEL). ProM es compatible con modelos para representar las reglas (por ejemplo, reglas basadas en LTL), las redes sociales y estructuras organizacionales.

<sup>5</sup> Un plugin es aquella aplicación que, en un programa informático, añade una funcionalidad adicional o una nueva característica al software.

Múltiples *plugins* están disponibles para cada una de las actividades que se muestran en el marco de trabajo antes expuesto, por ejemplo, hay docenas de *plugins* para descubrir y comprobar la conformidad de los modelos de procesos. Recientemente se han desarrollado *plugins* para apoyar la detección, predicción y recomendación de actividades. (Yzquierdo, 2013)

***Dotted chart analysis***: es un plugin en ProM, es una tabla similar a un diagrama de Gantt. Muestra una variedad de eventos de un registro de eventos a través del tiempo. La idea básica de la tabla de puntos es graficar puntos según el tiempo.

El *plugin* tiene un panel de menú de la izquierda, el gráfico se muestra en el medio y tanto la vista total del diagrama y las métricas explicadas aparecen a la derecha.

En la tabla, un punto en el gráfico representa un único evento en el registro, la tabla tiene dos dimensiones ortogonales: (1) vez y tipos de componentes (2). El tiempo se mide en el eje horizontal del gráfico. A lo largo del eje vertical, se muestran los tipos de componentes como instancia, creador, tarea, tipo de evento o elementos de datos.

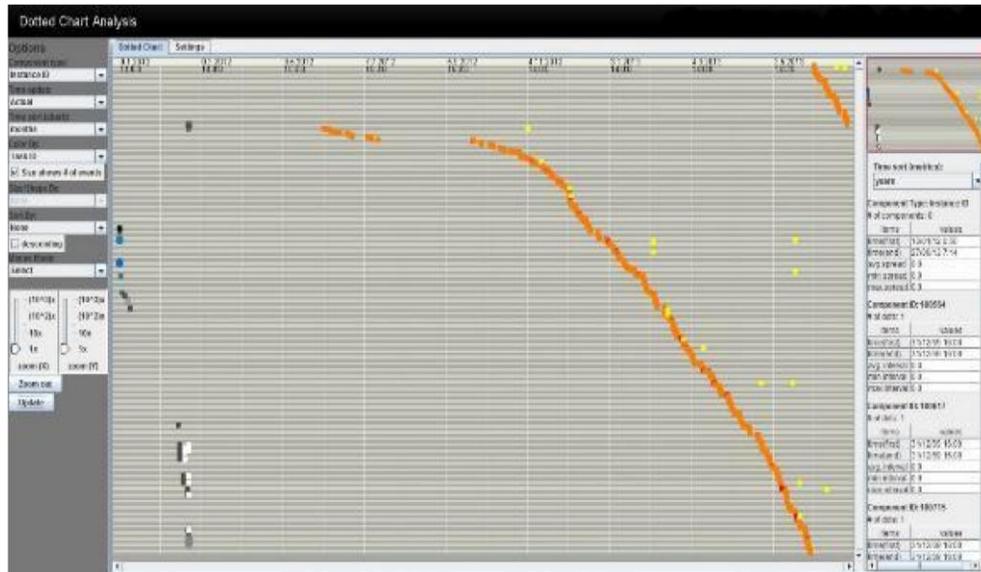
Proporciona las métricas relacionadas con eventos y su distribución en el tiempo. Hay dos tipos de indicadores de rendimiento: (1) métricas para el registro de eventos y (2) métricas para cada componente general. (Process Mining Group, Math&CS department, Eindhoven University of Technology, 2009)

Funcionalidades para el registro de eventos:

- La posición del primer evento en el registro.
- La posición del último evento en el registro.
- Promedio de propagación.
- Propagación mínima.
- Extensión máxima.

Funcionalidades para cada tipo de componente:

- La posición del primer evento de un componente.
- Puede calcularse la posición del último evento de un componente.
- Intervalo promedio entre eventos.
- Mínimo y máximo intervalo entre eventos.



**Figura 1.3:** Resultado de la técnica "Analyze using Dotted Chart".

Fuente: Process Mining Group, 2009.

- **Herramientas comerciales**

**PPMChart:** es creada bajo la suposición de que es posible grabar datos en cada operación de modelado (por ejemplo, crear evento de inicio, generar actividad, mover la actividad). Además del nombre de cada operación, la visualización necesita dos atributos más: el identificador del elemento modelo en que la operación fue realizada y la fecha y hora de la ejecución de la operación. Posiblemente otros atributos registrados (como la posición de un elemento modelo) son ignorados por la visualización.

El eje horizontal representa predeterminadamente un intervalo de tiempo de una hora. Verticalmente, cada línea representa un elemento del modelo de proceso, ya presentes en el modelado. El identificador del elemento modelo aparece al principio de la línea. Cada punto de color en la línea representa una operación realizada en el elemento.

- El color del punto se corresponde con el tipo de operación: crear (verde), mover (azul), borrar (rojo) y renombrar (naranja).
- La sombra del color y la forma del punto, se corresponden con el tipo de elemento modelo: actividad (brillante, caja), eventos (muy clara, círculo) y *gateway* (oscura, diamante)

- La posición del punto en la línea de tiempo se corresponde con el momento en que la operación fue ejecutada.

El usuario puede configurar el orden en que se presentan las líneas de tiempo. El orden predeterminado de líneas se corresponde con el orden lógico desde el evento de inicio hasta el evento final de los elementos en el modelo de proceso. (Jan Claes I. V., 2012)

**ARIS Process Performance Manager (ARIS PPM):** es una herramienta que permite a las empresas supervisar y analizar el rendimiento y la estructura de sus procesos empresariales. ARIS PPM optimiza continuamente los flujos de trabajo internos y externos, realizando de este modo una contribución crucial para el éxito de la empresa.

Esta herramienta muestra la información de forma gráfica utilizando tiempos de flujo y cuellos de botella, entre otros, y que sirven para realizar análisis en tiempo real. (Weske, van der Aalst, & Verbee, 2004)

La funcionalidad exclusiva de ARIS PPM permite a los usuarios visualizar automáticamente modelos de procesos reales basándose en datos del sistema de aplicación y de este modo, realizar comparaciones de los objetivos frente a los valores reales de los flujos de trabajo. (Manager, ARIS Process Performance, 2011)

Ventajas de ARIS PPM:

- El rendimiento de los procesos integrales resulta transparente.
- Medición del rendimiento de los procesos (cantidad, tiempo, coste, calidad).
- Supervisión perfecta de los valores de destino y del estado actual.
- Análisis de comunicación, actividades de departamentos y otras unidades organizativas.

### **Nivel nacional**

En Cuba no se encuentra documentación publicada que valide la existencia de algún sistema de minería de proceso que realice análisis de procesos desde la perspectiva temporal, solamente se han realizado investigaciones haciendo uso de herramientas como ProM y Disco, en distintos sectores de la sociedad.

## Resultados del análisis de las soluciones existentes

Tras concluir el análisis sobre los sistemas existentes para realizar Minería de procesos, que permitan efectuar análisis de procesos aplicando la perspectiva temporal, se constató que existen a nivel internacional una serie de herramientas implementadas con dicho propósito; las cuales tienen en común un conjunto de funcionalidades como son: mostrar la duración media de las actividades, la media de los tiempos de espera, la posición del primer y último evento en el registro, entre otras. A pesar de presentar tales características, las herramientas presentan limitantes, a continuación se muestran ejemplos de dichas limitantes:

- El plugin *Dotted chart analysis* de ProM no realiza análisis de procesos aplicando la perspectiva temporal, se limita a mostrar en un gráfico de puntos la ejecución temporal como se encuentra en el registro de eventos.
- ARIS PPM y PPM Chart están desarrolladas con fines comerciales y son privativas.
- Disco, PPM Chart y ARIS PPM son herramientas de escritorio lo que se hace imposible la integración con el HIS, al ser este una aplicación web.
- PPM Chart y ARIS PPM son herramientas que están dirigidas al sector industrial y empresarial.

Ninguna de las herramientas descritas soluciona la problemática de la investigación, debido a las limitantes identificadas, por lo que los autores deciden desarrollar un componente aplicando la perspectiva temporal de Minería de procesos a los registros de eventos obtenidos del HIS.

## 1.2. Tecnologías y herramientas

A continuación se describen las principales herramientas y tecnologías a utilizar en el proceso de desarrollo del componente propuesto, las cuales de son de software libre.

### 1.2.1. Marcos de trabajo

Los marcos de trabajo (*frameworks*) se pueden considerar como soluciones completas que contemplan herramientas de apoyo a la construcción (ambiente de trabajo o desarrollo) y motores de ejecución (ambiente de ejecución). Estos aceleran el proceso de desarrollo, permiten reutilizar código ya existente y promover buenas prácticas de desarrollo, como el uso de patrones. (Zaninotto, 2008)

### 1.2.2. Metodología de desarrollo

Las metodologías de desarrollo de software son un conjunto de procedimientos, técnicas y ayudas a la documentación para el desarrollo de productos software. (Carrillo, 2008)

La metodología de desarrollo AUP-UCI es empleada en los proyectos productivos de la UCI. Se basa en una variación de la metodología Proceso Unificado Ágil (AUP, por sus siglas en inglés) en unión con el modelo CMMI<sup>6</sup> para desarrollo (CMMI-DEV) versión 1.3. El cual constituye una guía para aplicar las mejores prácticas en una entidad desarrolladora. Estas prácticas se centran en el desarrollo de productos y servicios de calidad. (Chrissis, 2009) Esta metodología proporcionará la guía necesaria para el desarrollo del componente.

### 1.2.3. Lenguajes

#### Lenguaje Unificado de Modelado: *UML 2.1*

El Lenguaje Unificado de Modelado (UML, por sus siglas en inglés) es un lenguaje que permite modelar, construir y documentar los elementos que forman un sistema de software orientado a objetos. UML sirve para el modelado completo de sistemas complejos, tanto en el diseño de los sistemas de software como para la arquitectura hardware donde se ejecuten. (Orallo, 2012) Permite la representación conceptual y física de un sistema. Cuenta con varios tipos de diagramas, los cuales muestran diferentes aspectos de lo que se quiere representar. Ofrece un estándar para describir un plano del sistema, incluyendo aspectos conceptuales tales como procesos de negocios y funciones del sistema, y aspectos concretos como expresiones de lenguajes de programación, esquemas de bases de datos y componentes de software reutilizables. (Mora, 2013)

#### Lenguaje de programación: *Java 1.6*

El lenguaje de programación **Java** es robusto, multiplataforma. Tiene muchas similitudes con el lenguaje C y C++. La principal característica de Java es la de ser un lenguaje compilado e interpretado. Todo programa en Java ha de compilarse y el código que se genera es interpretado por una máquina virtual. ( The Apache Software Foundation, 2012) Permitirá la creación e implementación de las clases controladoras.

---

<sup>6</sup>Modelo de Madurez de Capacidad Integrado (CMMI, por sus siglas en inglés): es un modelo de madurez de mejora de los procesos para el desarrollo de productos y de servicios.

### **1.2.4. Tecnologías a utilizar**

Para el desarrollo del componente, a continuación se propone un conjunto de tecnologías y herramientas de código abierto, las cuales cumplen con las políticas de independencia tecnológica definidas en Cuba para la informatización de la sociedad cubana.

#### **Java Server Faces (JSF) 1.2**

Java Server Faces (JSF) es un marco de trabajo que define un modelo de componentes de interfaz de usuario y de eventos. Permite manejar el estado de los componentes de usuario, manejar sus eventos, la validación y conversión del lado del servidor y centralizar la navegabilidad de las páginas de la aplicación. JSF es el marco estándar que proporciona Java para construir aplicaciones web, sigue el patrón Modelo Vista Controlador (MVC), proporcionando una manera de validar datos, llamar a reglas de negocio, y devolver los resultados al cliente. (Introducción a JSF (Java Server Faces, 2012)

#### **Java Platform Enterprise Edition (JavaEE) 5.0**

Java versión 5 o JEE 5 es una plataforma de programación (parte de la Plataforma Java) para desarrollar y ejecutar software de aplicaciones en lenguaje de programación Java con arquitectura de N niveles distribuida (Franky, 2010). Se basa ampliamente en componentes de software modulares y se ejecuta sobre un servidor de aplicaciones.

#### **Java Persistence API (JPA)**

JPA proporciona un modelo de persistencia basado en la programación orientada a objeto para mapear bases de datos relacionales en Java. El *Java Persistence API* fue desarrollado por el grupo de expertos de EJB 3.0 como parte de JSR 220, aunque su uso no se limita a los componentes software EJB. También puede utilizarse directamente en aplicaciones web y aplicaciones clientes; incluso, fuera de la plataforma Java EE, por ejemplo, en aplicaciones Java SE. (Vázquez, 2006)

#### **Hibernate 3.3**

*Hibernate* es un marco de trabajo de persistencia para Java de libre distribución que facilita el mapeo de atributos entre una base de datos relacional y el modelo de objetos de una aplicación. Además proporciona un potente lenguaje de consultas denominado *Hibernate Query Language* (HQL). (Scribd., 2012)

### **SEAM 2.1**

SEAM es un marco de trabajo que integra la capa de presentación (JSF) con la capa de negocios y persistencia (EJB), funcionando, según versa su significado en español, como una “costura” entre estos componentes. Es una potente plataforma de desarrollo de código abierto para construir aplicaciones ricas de Internet en Java. *Seam* integra tecnologías como *JavaScript* asíncrono y XML (AJAX), *JavaServer Faces* (JSF), *Java Persistence Api* (JPa), *Enterprise Java Beans* (EJB 3.0) y *Business Process Management* (BPM). (Álvarez, Mora, Hernández, & Alemán, 2013)

### **Java Runtime Environment (JRE)**

*JRE* es el acrónimo de *Java Runtime Environment* (entorno en tiempo de ejecución Java) y se corresponde con un conjunto de utilidades que permite la ejecución de programas java sobre todas las plataformas soportadas. JVM (máquina virtual Java) es una instancia de JRE en tiempo de ejecución. Este interpreta el código Java y está compuesto además por las librerías de clases estándar que implementan el API de *Java*. Ambas JVM y API deben ser consistentes entre sí, de ahí que sean distribuidas de modo conjunto. (Lucifer, n.d.)

### **Jboss Server 4.2.2**

*Jboss Server* es un servidor de aplicaciones J2EE de código abierto implementado en Java puro. Al estar basado en Java, *JBoss* puede ser utilizado en cualquier sistema operativo para el que esté disponible Java. Proporciona una herramienta útil para el desarrollo y despliegue de aplicaciones Java, aplicaciones Web y portales. *JBoss AS* puede ser descargado, utilizado, incrustado y distribuido sin restricciones por la licencia. (Consejería de Hacienda y Administración Pública, 2007) Permitirá la publicación de las páginas XHTML.

### **Facelets 1.1**

Facelets es un *framework* simplificado de presentación, donde es posible diseñar de forma libre una página web y luego asociarle los componentes JSF específicos. Aporta mayor libertad al diseñador y mejora los informes de errores que tiene JSF. (Sánchez, 2008 )

### **RichFaces 3.3.1**

Es una biblioteca de componentes web enriquecidos, de código abierto y basada en el estándar *Java Server Faces* (JSF). Provee facilidades de validación y conversión de los datos

proporcionados por el usuario, administración avanzada de recursos como imágenes, código *Javascript* y Hojas de Estilo en Cascada (CSS). (Gaguancela, 2012).

Permitirá crear interfaces de usuario modernas de manera rápida, basadas en los componentes listos para usar, altamente configurables en cuanto a temas predefinidos por el propio marco de trabajo o desarrollados a conveniencia. Específicamente en el componente a desarrollar, *richfaces* permitirá la visualización de la información.

### ***Enterprise JavaBeans (EJB) 3.0***

Es un componente utilizado en Java que permite agrupar funcionalidades para formar parte de una aplicación, esto puede ser: un "*Java Bean*" agrupando información personal, datos sobre un pedimento, requerimientos de órdenes, entre otros. Permite realizar la administración automática de transacciones, seguridad, escalabilidad, concurrencia, distribución, acceso a ambientes portables y persistencia de datos. Incorpora el estándar JPA como el principal API de persistencia para aplicaciones EJB3. Su objetivo es simplificar el desarrollo de aplicaciones Java y estandarizar el API de persistencia para la plataforma *Java*. Forma parte de la especificación JEE 5. (Rondón, 2009)

### **1.2.5. Herramientas a utilizar**

#### ***Herramienta CASE Visual Paradigm 6.4***

Las herramientas CASE son un conjunto de programas y ayudas que dan asistencia a los analistas, ingenieros de software y desarrolladores, durante todos los pasos del ciclo de vida de desarrollo de un Software. (Alfaro, 2011)

Visual Paradigm 6.4 es una herramienta de UML para el desarrollo de software de aplicación, diseñada para la ayuda del desarrollo del software. Es utilizada por distintos usuarios entre los que se incluyen ingenieros de software, analistas de sistemas, analistas de negocios, arquitectos y desarrolladores. Soporta el ciclo de vida completo del desarrollo de software. (Santiesteban, et al., 2011) Esta herramienta permitió la modelación del componente utilizando como lenguaje UML a través de la creación de diagramas en un ambiente visual.

### ***Entorno Integrado de Desarrollo Eclipse 3.4.2***

Un Entorno de Desarrollo Integrado (IDE, por sus siglas en inglés) es un entorno de programación que ha sido empaquetado como un programa de aplicación, es decir, consiste en un editor de código, un compilador, un depurador y un constructor de interfaz gráfica (GUI). Los *IDEs* pueden ser aplicaciones por sí solas o pueden ser parte de aplicaciones existentes. Proveen un marco de trabajo amigable para la mayoría de los lenguajes de programación. (Alfaro, 2011)

Es Eclipse *Ganymede* principalmente una plataforma de programación, usada para crear entornos integrados de desarrollo. Pese a que Eclipse esté escrito en su mayor parte en Java y su uso más popular sea como un IDE para Java, Eclipse es neutral y adaptable a cualquier tipo de lenguaje de programación. Es una gran estructura formada por un núcleo y muchos *plugins* que van conformando la funcionalidad final. (Varga, 2007) Para el desarrollo del componente se le integró el marco de trabajo *JBoss Seam*. Soportará el lenguaje de programación Java para la implementación de las clases.

### ***Sistema Gestor de Base de Datos (SGBD)***

Un sistema gestor de base de datos se define como el conjunto de programas que administran y gestionan la información contenida en una base de datos. (Álvarez, 2007)

Como Sistema Gestor de Base de Datos (SGBD) se utilizó PostgreSQL 9.3.3 que es un conjunto de programas no visibles al usuario final que se encargan de la privacidad, la integridad, la seguridad de los datos y la interacción con el sistema operativo. (Oquendo, 2011) La versión 9.3.3 es la última estable del producto, que fue desarrollado por una comunidad conocida como PostgreSQL *Global Development Group* (PGDG). Por las características antes mencionadas y por ser de amplia utilización en nuestra universidad, se decidió utilizar este sistema.

### ***Plugin Add Identities to log de ProM***

El plugin *Add identities to Log* de ProM muestra un resumen de toda la información contenida en el registro de eventos, a continuación se presenta una descripción funcional de dicho plugin:

- Número total de instancias de procesos: muestra el total de instancias de procesos que se registraron en el registro de eventos.

- Número total de eventos: muestra el total de eventos que se registraron en el registro de eventos.
- Todos los eventos: muestra los nombres de los eventos que se registran en el registro de eventos, además de la cantidad de ocurrencias absoluta y la cantidad de ocurrencias relativas.
- Eventos de inicio: muestra los eventos que inician cada instancia del proceso.
- Eventos de fin: muestra todos los eventos que finalizan cada instancia del proceso.
- Recurso: muestra el nombre de los recursos que ejecutan los eventos registrados en el registro de eventos.

El plugin *Add identities to Log* de ProM se utilizará para la validación de la propuesta de solución.

### **Conclusiones parciales**

El análisis crítico y valorativo a los sistemas existentes de Minería de procesos aplicando la perspectiva temporal, permitió comprobar la existencia de herramientas que realizan análisis desde este punto de vista. Sin embargo, estas herramientas presentan limitantes que impiden su utilización en la presente investigación. Por tanto, los autores deciden desarrollar un componente aplicando la perspectiva temporal de Minería de procesos a los registros de eventos obtenidos del HIS. El análisis realizado a las principales técnicas de descubrimiento, permitió seleccionar la técnica Minería de variantes, por su capacidad de generar modelos de variantes de procesos en forma de árbol.

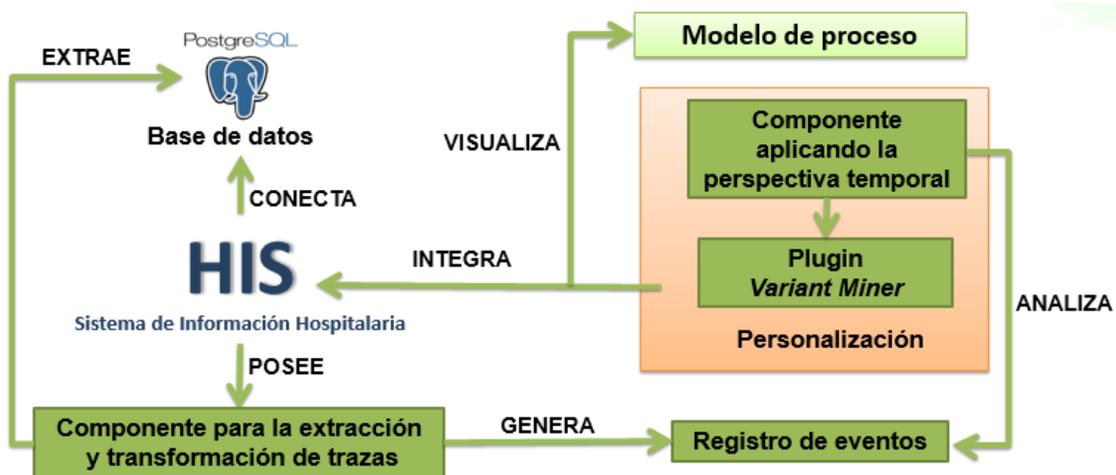
## CAPÍTULO 2: PROPUESTA DE SOLUCIÓN.

El objetivo fundamental del capítulo es describir la propuesta de solución, abordar los aspectos principales de la realización del Componente para la visualización de modelos de procesos aplicando la perspectiva temporal, describir el flujo información y definir la arquitectura con la que cuenta el software y los patrones de diseños tenidos en cuenta para la realización del componente y mostrar el procedimiento seguido para el desarrollo del mismo.

### 2.1. Propuesta de solución

La presente investigación propone desarrollar un componente para la visualización de modelos de procesos aplicando la perspectiva temporal. Para el desarrollo del mismo se sigue el siguiente flujo de información.

El componente para la extracción y transformación de trazas del HIS, extrae la información de la ejecución de las actividades de una base de datos y genera un registro de eventos, el cual es el punto de partida del componente aplicando la perspectiva temporal. Luego se personaliza el plugin *Variants Miner* incorporándole el componente antes mencionado. La personalización se integra al HIS y genera un modelo de proceso, el cual se visualiza desde el sistema. En la **Fig. 2.1**, se muestra el flujo de información de la propuesta de solución.



**Figura 2.1:** Flujo de información de la propuesta de solución

Fuente: Elaboración propia.

Para desarrollar el componente propuesto se identificó el requisito Generar modelo de proceso. A continuación se relaciona los productos de trabajo de ingeniería elaborados para la implementación del requisito:

- Especificación de requisitos de software: ([ver anexo 1](#))
- Especificación de caso uso: ([ver anexo 2](#))
- Diseño de casos de prueba: ([ver anexo 3](#))

### 2.2. Patrones arquitectónicos

Los patrones arquitectónicos, o patrones de arquitectura, son patrones de diseño de software que ofrecen soluciones a problemas de arquitectura de software en ingeniería de software. Dan una descripción de los elementos y el tipo de relación que tienen junto con un conjunto de restricciones sobre cómo pueden ser usados. Un patrón arquitectónico expresa un esquema de organización estructural esencial para un sistema de software, que consta de subsistemas, sus responsabilidades e interrelaciones. (Ivar Jacobson, 2009)

Para el desarrollo del componente para la visualización de modelos de procesos aplicando la perspectiva temporal de Minería de procesos, se utiliza el patrón arquitectónico Modelo Vista Controlador (MVC).

El patrón conocido como Modelo Vista Controlador (MVC) separa el modelado del dominio, la presentación y las acciones basadas en datos ingresados por el usuario en tres capas diferentes, como se explicará a continuación.

**Modelo:** el modelo administra el comportamiento y los datos del dominio de la aplicación, responde a requerimientos de información sobre su estado y responde a instrucciones de cambiar el estado.

**Vista:** maneja la visualización de la información. Están contenidas las páginas cliente, bibliotecas y marcos de trabajo encargados de esta tarea.

**Controlador:** interpreta las acciones del usuario informando al modelo y/o a la vista para que cambien según resulte apropiado.

Tanto la vista como el controlador dependen del modelo, el cual no depende de las otras capas. Esta separación permite construir y probar el modelo independientemente de la representación visual.

### 2.3. Patrones de diseño

Un patrón es una descripción, que se codifica en un formato estructurado, de un problema y su solución que recibe un nombre y que puede emplearse en otros contextos; en teoría, indica la manera de utilizarlo en circunstancias diversas. Muchos patrones ofrecen orientación sobre cómo asignar las responsabilidades a los objetos ante determinadas situaciones.

#### 2.3.1. Patrón de diseño GRASP

Los Patrones de Software de Asignación de Responsabilidades (GRASP por sus siglas en inglés), son aquellos que describen los principios fundamentales de la asignación de responsabilidades a objetos. El propósito de los mismos de forma general es originar componentes robustos, entendibles y fáciles de mantener y reutilizar, lo cual explica que su adecuada utilización sea la clave para un exitoso diseño.

A continuación se describen cuatro patrones GRASP que se tuvieron en cuenta para el diseño.

**Patrón experto:** es el principio básico de asignación de responsabilidades. Nos indica, por ejemplo, que la responsabilidad de la creación de un objeto o la implementación de un método, debe recaer sobre la clase que conoce toda la información necesaria para crearlo. De este modo obtendremos un diseño con mayor cohesión y así la información se mantiene encapsulada (disminución del acoplamiento). (Almaguer, 2011)

Se evidencia en la clase controladora `VariantsTreeManager`, la cual, tiene la responsabilidad de ejecutar la descomposición en subprocesos a partir de los parámetros de entrada y del registro de evento.

**Patrón creador:** el patrón creador nos ayuda a identificar quién debe ser el responsable de la creación (o instanciación) de nuevos objetos o clases. La nueva instancia deberá ser creada por la clase que: tiene la información necesaria para realizar la creación del objeto, usa directamente las instancias creadas del objeto, almacena o maneja varias instancias de la clase o contiene o agrega la clase. Una ventaja es el bajo acoplamiento, lo cual supone facilidad de mantenimiento y reutilización. (Almaguer, 2011)

Este patrón se evidencia en la clase controladora `VariantsTreePlugin`, encargada de la creación o instanciación de los objetos necesarios para el trabajo del mismo.

**Patrón alta cohesión:** los conceptos de cohesión y acoplamiento están íntimamente relacionados. Un mayor grado de cohesión implica un menor de acoplamiento. Plantea que la información que almacena una clase debe de ser coherente y debe estar (en la medida de lo posible) relacionada con la clase. (Almaguer, 2011)

La totalidad de las clases del componente evidencian el uso de este patrón, cada una contiene, principalmente, la información que le pertenece de forma coherente.

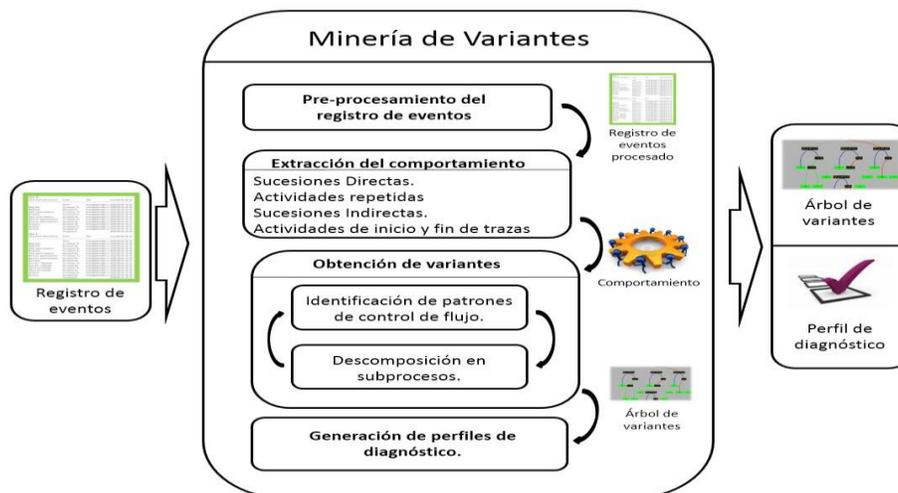
**Patrón Bajo Acoplamiento:** es la idea de tener las clases lo menos ligadas entre sí que se pueda. De tal forma que en caso de producirse una modificación en alguna de ellas, se tenga la mínima repercusión posible en el resto de clases, potenciando la reutilización, y disminuyendo la dependencia entre las clases. (Almaguer, 2011) Teniendo en cuenta que se implementaron clases altamente cohesionadas se garantizó el bajo acoplamiento de las mismas.

### 2.4. Técnica de Minería de variantes

La Minería de variantes está conformada por las siguientes etapas, las cuales se ejecutan secuencialmente:

1. Pre-procesamiento del registro de eventos.
2. Extracción de comportamiento.
3. Búsqueda de variantes.

Existe una cuarta etapa que no es de interés para la presente investigación.



**Figura 2.2:** Esquema representativo de la técnica Minería de variantes.

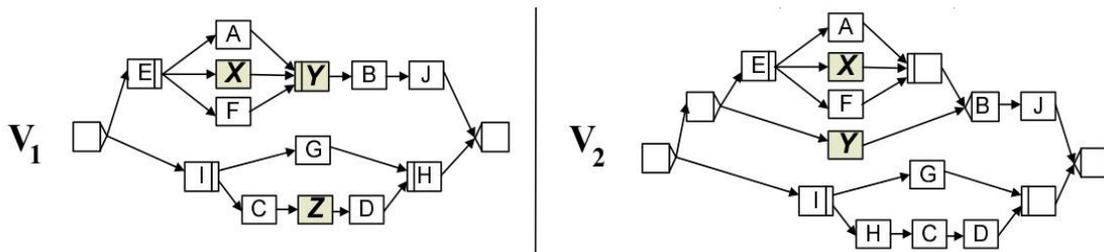
Fuente: Alfonso, 2014.

Para comprender el árbol de variantes que ofrece la técnica como resultado, es relevante conocer que la descomposición que se realiza en la técnica se basa en la siguiente definición:

**Descomposición en subprocesos:** Se denota con  $S_P$  al conjunto de los subprocesos del proceso  $P$ , y por  $W = \{ws, wx, wo, wp, wl\}$  se denota al conjunto de los patrones de control de flujo. Un subproceso  $S_i \in S_P$  se puede descomponer a través de diferentes  $w_j | \forall w_j \in W$  en varios  $s_k | \forall k \neq i : s_k \in S_P$ , hasta el nivel de actividad.

La técnica propone varias descomposiciones alternativas para el mismo subproceso, utilizando diferentes patrones de control de flujo. Esto permite controlar el impacto estructural del ruido y la ausencia de información en la construcción de las alternativas. Las alternativas se construyen descartando o no, determinados comportamientos presentes en el registro de eventos. También se pueden construir, considerando o no, determinados comportamientos ausentes. Las diferentes alternativas de descomposición que pueden existir en cada subproceso conforman variantes del proceso.

**Variantes de proceso:** Las variantes de un modelo de proceso o variantes de proceso, son modelos de un proceso que describen el mismo proceso de negocio, y poseen algunas diferencias estructurales. Las diferencias están dadas por los patrones de control de flujo que se utilizan en secciones equivalentes del proceso y la presencia de determinadas actividades.



**Figura 2.2:** Variantes de proceso.

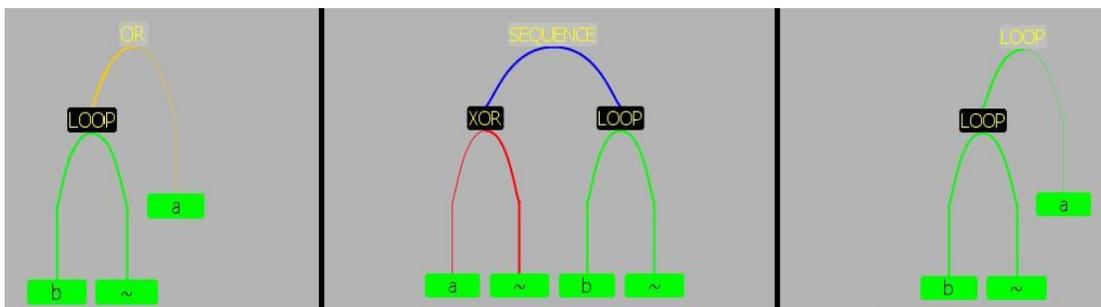
Fuente: Alfonso, 2014.

En la **Fig. 2.2**, se aprecian dos variantes (V1 y V2) de un proceso de negocio. En V1, la actividad Y se muestra en secuencia con el subproceso conformado por las actividades E, A, X y F; por

otro lado en V2, la actividad Y aparece en selección exclusiva con el subproceso antes mencionado. Además la actividad Z, aparece en V1, pero no en V2.

Los elementos antes planteados permiten que la técnica ofrezca un modelo jerárquico en forma de árbol, que representa diferentes variantes del proceso cuya ejecución está contenida en el registro de eventos. Para esto se propone un modelo de proceso denominado árbol de variantes.

En la **Fig. 2.3** se muestran las variantes de descomposición para un mismo subproceso, que forman parte de un árbol de variantes.



**Figura 2.3:** Variantes contenidas en un árbol de variantes.

Fuente: Alfonso, 2014.

La primera variante corresponde al patrón selección no exclusiva. La segunda variante representa una descomposición por el patrón de control de flujo secuencia. Por último, se muestra la descomposición del subproceso utilizando el patrón lazo. (Alfonso, 2014)

### 2.4.1. Etapas de la Minería de variantes

A continuación se definen las etapas por las cuales está compuesta la técnica de Minería de variantes.

#### Pre-procesamiento del registro de eventos

Esta etapa tiene como objetivo preparar el registro de eventos para el diagnóstico de variantes. El algoritmo que se ejecuta en esta etapa realiza la codificación de los eventos presentes en el registro de eventos. La codificación consiste en la asignación de una letra a cada evento contenido en el registro. Esto tiene como objetivo disminuir la carga cognitiva durante la comprensión del proceso, eliminando la información semántica de las etiquetas de los eventos. Mediante la codificación se pretende mejorar la comprensión sintáctica del modelo, antes de

pasar a la comprensión semántica del mismo. Adicionalmente, la conversión de las trazas en secuencias codificadas facilita el procesamiento subsecuente.

Posterior a la codificación se unifican las trazas cuyas secuencias coinciden, disminuyendo el número de trazas a procesar y agrupando los comportamientos coincidentes. Las secuencias resultantes de la agrupación son ordenadas descendientemente, por su frecuencia de aparición, para expresar su importancia relativa dentro del proceso. Por último se extraen características generales presentes en el registro de eventos, como la cantidad de trazas diferentes y la cantidad de eventos diferentes.

### Extracción de comportamiento

El propósito de esta etapa es la extracción de comportamientos representativos de los patrones de control de flujo, a partir del registro de eventos pre-procesado. Para la técnica se considera el comportamiento del proceso: se denota por  $l_i \subseteq L_p$  a la sección del registro de eventos relacionada con el subproceso  $s_i \in S_p$ . El comportamiento del patrón de control de flujo  $w_j$  para el subproceso  $s_i$ , denotado por  $\beta_{w_j s_i} | w_j \in W$  está compuesto por todas las posibles:

- Sucesiones directas (*DS*), se denomina sucesión directa entre la actividad a y la actividad b, a la existencia en alguna traza de la secuencia ab.
- Sucesiones indirectas (*IS*), se denomina sucesión indirecta entre la actividad a y la actividad b, a la aparición en alguna traza de la actividad a seguida, inmediatamente o no, por la actividad b.
- Eventos que inician y/o finalizan trazas.
- Eventos repetidos en las trazas.

Siendo  $l_i$  la sección del registro de eventos relacionada con el subproceso  $s_i$ , el subconjunto de este comportamiento observable en  $l_i$  es el comportamiento del proceso  $w_j$  en  $l_i$ :  $\beta_{w_j l_i} \subseteq \beta_{w_j s_i} | w_j \in W$ . Por tanto, el comportamiento presente en la sección del registro de eventos  $l_i$  es  $\beta_{l_i} = \beta_{w_s l_i} \cup \beta_{w_x l_i} \cup \beta_{w_o l_i} \cup \beta_{w_p l_i} \cup \beta_{w_l l_i}$ .

Por cada comportamiento extraído se obtiene la frecuencia de su aparición en las trazas codificadas. Esta frecuencia es utilizada para establecer cuán representativo es un comportamiento del proceso que lo genera.

Los comportamientos extraídos en esta fase se combinan durante la obtención de variantes de descomposición. La extracción de comportamiento permite identificar, con mayor precisión que a nivel de traza, los elementos de ejecución del proceso, cuya baja frecuencia de ejecución sugiere que deban ser considerados como ruido.

### **Obtención de variantes de descomposición**

En la tercera etapa se obtienen las variantes de descomposición del proceso para su representación en un árbol de variantes. Las variantes se obtienen a partir de la combinación de los comportamientos extraídos en la fase anterior y la asunción de ciertos comportamientos ausentes del registro de eventos. En esta fase se ejecutan dos algoritmos iterativamente, uno para la descomposición en subprocesos y otro para la identificación de patrones de control de flujo.

### **2.5. Desarrollo del componente para la visualización de modelos de procesos aplicando la perspectiva temporal de Minería de procesos**

El componente para la visualización de modelos de procesos aplicando la perspectiva temporal de Minería de procesos, utiliza como entrada un registro de eventos en formato .XES y genera como salida un modelo de procesos en forma de Árbol, donde cada actividad mostrada en el árbol contendrá su duración máxima, media, mínima y total.

A continuación se muestra el procedimiento descrito en tres fases, para el desarrollo de dicho componente.

1. El desarrollo de un componente aplicando la perspectiva temporal de Minería de Procesos.
2. La personalización del plugin *Variants Miner*, incorporándole el componente desarrollado y determinando las configuraciones para el HIS.
3. La integración de la personalización del plugin *Variants Miner* en el HIS del CESIM, para la visualización de modelos de procesos.

### 2.5.1. Componente aplicando la perspectiva temporal de Minería de procesos

El componente aplicando la perspectiva temporal Minería de procesos se desarrolla con el objetivo de obtener las métricas para el análisis de procesos aplicando la perspectiva temporal, dichas métricas son: la duración mínima, media, máxima y total de cada una de las actividades del proceso que se está analizando.

**Duración total:** se refiere a la suma de todos los tiempos de duración de las instancias de cada actividad de un proceso.

**Duración media:** es el tiempo promedio en que se ejecutan las actividades de un proceso, la misma se realiza basándose en la siguiente fórmula de media aritmética:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n a_i = \frac{a_1 + a_2 + \dots + a_n}{n}$$

**Figura 2.4:** Fórmula de media aritmética.

Fuente: Revista de Didáctica de las Matemáticas, 1995 JUL; II (5).

Una actividad dentro de un proceso puede ejecutarse una o varias veces, cada  $a_i$  representa el tiempo que dura la actividad en la aparición  $i$  y  $n$  el total de apariciones de dicha actividad.

**Duración mínima:** del total de apariciones, la menor duración registrada de la actividad.

**Duración máxima:** del total de apariciones, la mayor duración registrada de la actividad.

Para el desarrollo del componente, se crearon las clases `Analisis`, `TimeDifDate`, `Metriccs`.

La clase `TimeDifDate` es una estructura donde se almacena la diferencia de tiempo entre el inicio y el fin de una actividad.

La clase `Metriccs` representa una estructura para almacenar duración media, máxima, mínima y total de ejecución de las actividades en el registro de eventos.

La clase `Analisis` es una clase controladora, que permite calcular la duración media, la duración máxima y la duración mínima de cada una de las actividades, haciendo uso de las clases `TimeDifDate` y `Metriccs`.

### 2.5.2. Personalización del plugin *Variants Miner*

La personalización del plugin *Variants Miner* se realizó con el objetivo de:

1. Establecer las configuraciones para el HIS.
2. Mostrar en los nodos hojas del árbol que devuelve el plugin, las métricas generadas por el componente aplicando la perspectiva temporal.

#### **Objetivo 1: Establecer las configuraciones para el HIS**

La personalización al igual que el complemento de ProM, utiliza como entrada un registro de eventos en formato .XES y un ámbito definido como parámetros de entrada, el ámbito se especifica en términos de umbrales para ruido y completitud.

Los umbrales para ruido y completitud se definen como un número entre 0 y 100 y pueden ser especificados individualmente para cada patrón de control de flujo. Los valores serán asignados por defecto, donde los valores de los umbrales para ruido serán de un 5% y para los umbrales de completitud serán de un 95%, se establecieron dichos valores con el propósito de obtener un modelo que sea comprensible por el usuario y a la vez que brinde la mayor exactitud posible.

#### **Objetivo 2: Mostrar en los nodos hojas del árbol que devuelve el plugin, las métricas generadas por el componente aplicando la perspectiva temporal**

Para cumplir con el objetivo se creó un paquete denominado `org.processmining.variantstree.patterns.temporal`, el cual contiene las clases desarrolladas en el componente aplicando la perspectiva temporal de Minería de procesos.

### 2.5.3. Integración de la personalización del plugin *Variants Miner* en el HIS del CESIM, para la visualización de modelos de procesos

#### **Paso 1: Integración**

Una vez personalizado el plugin *Variants Miner*, se procede a su integración con el HIS. A continuación se detallan los pasos para realizar la integración:

1. El registro de eventos es utilizado como parámetro de entrada de la clase `AnalisisVariants`.

2. Luego se instancia la clase `VariantsTreeManager` que descompone el proceso almacenado en el registro de eventos, en varios subprocesos.
3. Posteriormente se obtiene el modelo, el cual se convierte a un archivo en formato `svg`, a partir del método `exportarSVG` de la clase `AnalisisVariants`.
4. Finalmente se visualiza en el HIS el modelo de proceso.

### Paso 2: Visualización

El componente incluye dentro de sus elementos fundamentales, la interfaz de usuario, la cual se estructura de las siguientes secciones:

- Área para la entrada de datos
- Área de resultados

A continuación se describen cada una de las secciones antes mencionadas.

- **Área para la entrada de datos**

Esta sección constituye el área donde los analistas e investigadores de procesos introducirán los datos relacionados que serán necesarios para que el sistema muestre el modelo, los cuales son: selección del proceso, selección del tipo de análisis, en este caso Perspectiva temporal y el rango de tiempo en el cual se obtendrán las trazas de ejecución para el proceso seleccionado.

Parámetros de entrada

Seleccionar proceso: proceso solicitar productos

Desde: 2011-11-10

Hasta: 2012-03-15

Seleccionar tipo de análisis: Perspectiva Temporal

Generar Cancelar

**Figura 2.5:** Área para la entrada de datos.

Fuente: Elaboración propia.

Para obtener una vista del modelo, se presiona el botón Generar.

- **Área de resultados**

En esta sección, el sistema muestra los resultados, el cual es un modelo de variantes de procesos en forma de árbol. Donde los nodos hojas representan las actividades que conforman el proceso, cada actividad aparece con su duración mínima, media, máxima y total. Los nodos patrón, muestran el nombre del patrón de control de flujo utilizado para descomponer el subproceso en la variante correspondiente, las aristas del árbol representan caminos dentro del proceso. El grosor de estas aristas depende de la frecuencia con la cual aparece en el registro de eventos el camino que representan y dicho grosor representa además las posibles desviaciones<sup>7</sup>, es decir mientras más finas sean las aristas, mayor existencia de desviaciones.

En el área también se muestra una leyenda especificando cada actividad del proceso, pues estas aparecen codificadas en el modelo, además en dicha leyenda se muestra la frecuencia<sup>8</sup> de cada una de las actividades. El sistema muestra los patrones de control de flujo utilizados para la descomposición en subproceso y si se selecciona uno de dichos patrones, el sistema muestra una descripción de su significado. El sistema permite expandir, contraer y arrastrar el modelo.

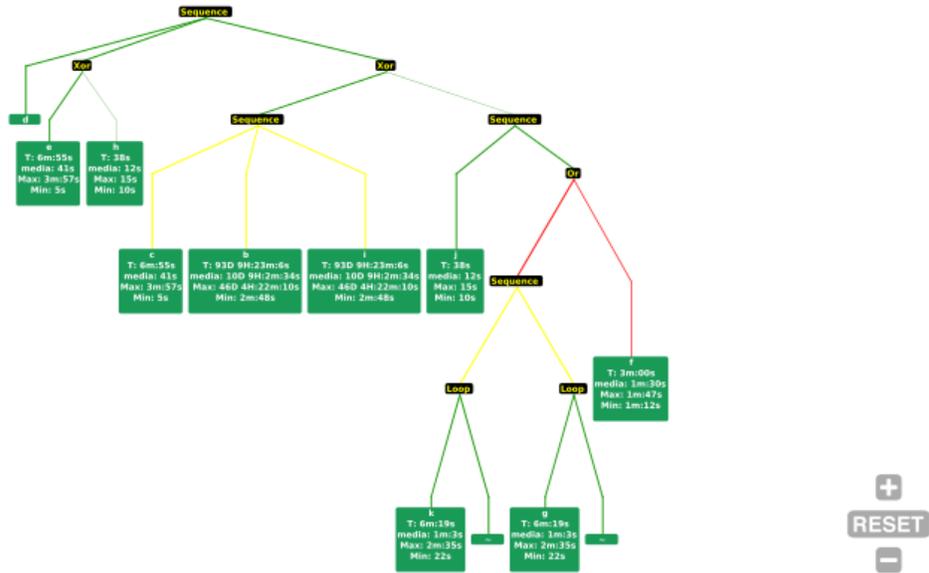
A continuación se muestra en la **Fig.2.6** una vista de dicha área desde el HIS.

---

<sup>7</sup> **Las desviaciones** muestran precisamente las partes del modelo que se desvían con respecto al registro de eventos, son visualizadas para mostrar que partes del modelo se ajustan bien y que partes no lo hacen.

**La frecuencia:** está dada por la cantidad de veces que se repite una actividad con respecto a todas las demás en un registro de eventos.

Perspectiva Temporal



Leyenda de actividades		
Código	Actividad	Frecuencia
f	modificar_solicitud_licitacion+complete	2
g	aceptar_niveles+complete	6
d	START+complete (actividad de inicio de traza, sin valor para el modelo)	13
e	ver_detalle_sol_bq+start	10
b	despacho_sol_bq+start	9
c	ver_detalle_sol_bq+complete	10
a	modificar_solicitud_licitacion+start	2
j	ver_detalle_sol_licitacion+complete	3
k	aceptar_niveles+start	6
h	ver_detalle_sol_licitacion+start	3
i	despacho_sol_bq+complete	9

PATRONES

- SEQUENCE (Secuencia)
- XOR (Selección Exclusiva)
- OR (Selección No Exclusiva)
- PARALLELISM (Paralelismo)
- LOOP (Lazo)

Figura 2.6: Área de resultados. Interfaz del “modelo” generado por el componente

Fuente: Elaboración propia

Conclusiones parciales

El desarrollo del componente aplicando la perspectiva temporal posibilitó la obtención de las métricas para el análisis de procesos. La personalización del plugin *Variants Miner* permitió mostrar en los nodos hojas del árbol de variantes las métricas obtenidas por el componente antes mencionado. La descripción del procedimiento realizado para el desarrollo del componente propuesto permitió detallar cada una de las fases ejecutadas para la realización del mismo.

### **CAPÍTULO 3: VALIDACIÓN DE LA SOLUCIÓN PROPUESTA**

El objetivo del presente capítulo es validar la solución propuesta y mostrar los resultados arrojados por el componente análisis de procesos aplicando la perspectiva temporal, a continuación se presentan los temas a tratar en el capítulo.

Temas a tratar en el presente capítulo:

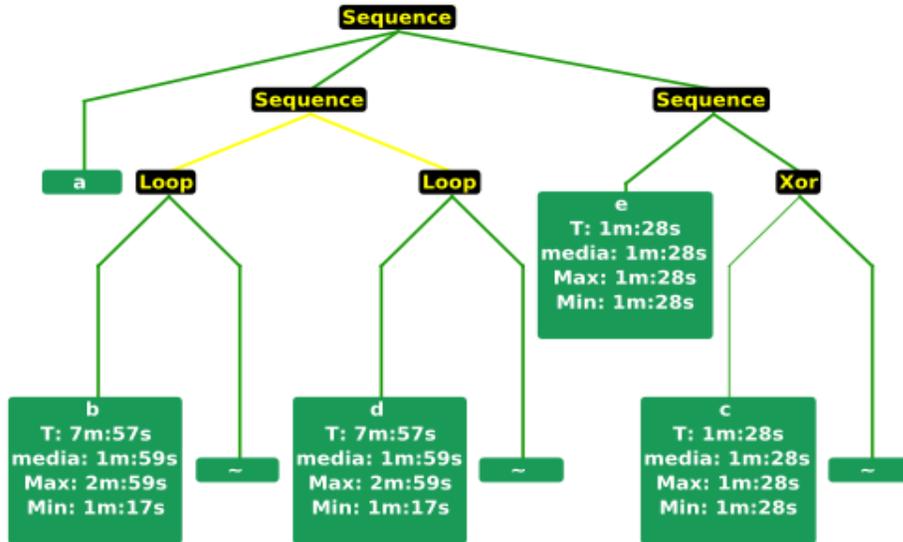
- Comprensión del modelo generado por el componente propuesto en el HIS
- Identificación de métricas para el análisis de procesos aplicando la perspectiva temporal
- Validación de las actividades

#### **3.1. Comprensión del modelo generado por el componente desarrollado en el HIS**

Para comprender los modelos generados por el componente en el sistema se ha aplicado el mismo, a un registro de eventos que corresponde al módulo Almacén del HIS. El registro de eventos recopila 3 ejecuciones (instancias) del proceso Procesar Solicitudes entre las fechas 10/05/2011 y 31/05/2012. Dicho proceso cuenta con 5 actividades, con un total de 15 ocurrencias.

A continuación se muestra el modelo generado:

Perspectiva Temporal



Leyenda de actividades	
Código	Actividad
d	crear_contrato+complete
e	crear_cronograma_de_entrega+start
b	crear_contrato+start
c	crear_cronograma_de_entrega+complete
a	START+complete (actividad de inicio de traza, sin valor para el modelo)

Figura 3.1: Modelo del proceso Procesar Solicitudes

Fuente: Elaboración propia

Como se observa en el modelo de la Fig. 3.1, el primer nivel de abstracción del árbol lo constituye el patrón de control de flujo Secuencia, lo que significa que los subprocesos *crear\_contrato* y *crear\_cronograma\_de\_entrega*, se ejecutaran uno seguido del otro.

Seguidamente se realiza el análisis de los subprocesos identificados correspondientes al segundo nivel de abstracción.

**Subproceso crear\_cronograma\_de\_entrega**

La descomposición obtenida para este subproceso y la codificación de sus actividades se muestran en la Fig. 3.2.

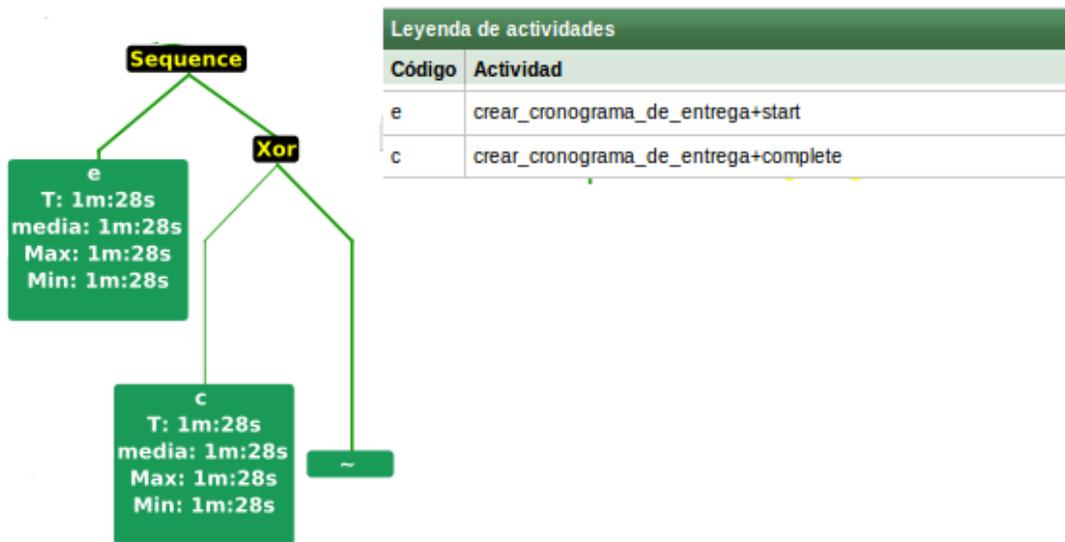


Figura 3.2: Modelo del subproceso crear\_cronograma\_de\_entrega

Fuente: Elaboración propia

En esta sección del registro de eventos se identifica la existencia de la secuencia entre la actividad *crear\_cronograma\_de\_entrega+star* y un subproceso descompuesto mediante el patrón de control de flujo selección exclusiva. La traza que contiene únicamente a la actividad *crear\_cronograma\_de\_entrega+complete* representa una de las situaciones de ausencia de información reportadas en la literatura. (Yzquierdo, 2012)

La ausencia de representación de uno de los subprocesos que se encuentran en secuencia en alguna traza es considerada ausencia de información para el patrón de control de flujo secuencia. El análisis de la sección del registro de eventos correspondiente evidencia la ausencia de información en la segunda traza.

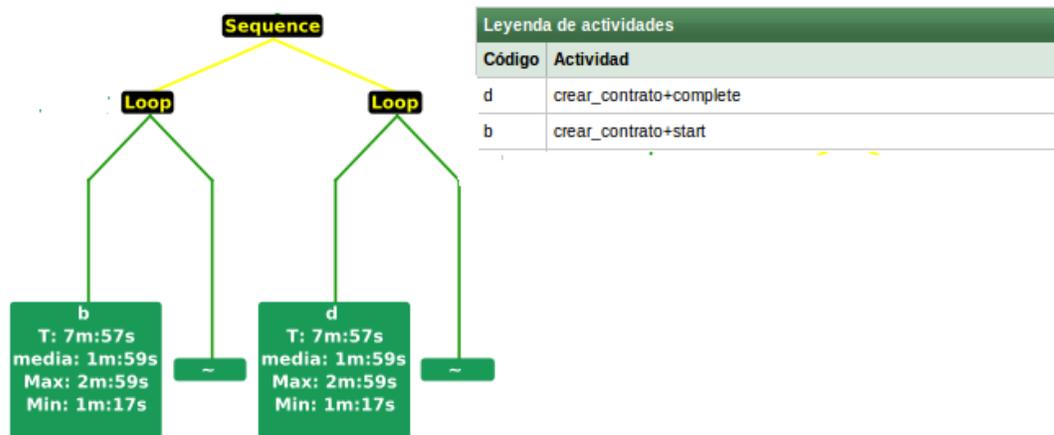
La identificación de la ausencia de información en esa traza provoca la inserción de una actividad invisible inmediatamente después de la actividad *crear\_cronograma\_de\_entrega+complete*. Una tarea invisible se puede manifestar cuando se produce un salto de una o varias actividades en una situación de selección. La inserción de la actividad invisible en la traza posibilita la posterior identificación del patrón de selección exclusiva.

En el árbol se puede apreciar en el grosor de las aristas que se refieren al subproceso cuya descomposición se realizó utilizando el patrón de selección exclusiva, que la alternativa que contiene la actividad invisible insertada es menos frecuente. Esto indica que la ejecución más

común del subproceso *crear\_cronograma\_de\_entrega* es la secuencia conformada por *crear\_cronograma\_de\_entrega+start* y *crear\_cronograma\_de\_entrega+complete*.

### Subproceso *crear\_contrato*

La descomposición obtenida para este subproceso y la codificación de sus actividades se muestran en la **Fig. 3.3**.



**Figura 3.3:** Modelo del subproceso *crear\_contrato*

Fuente: Elaboración propia

En esta sección del registro de eventos se identifica la existencia de la secuencia entre dos subproceso descompuesto mediante el patrón de control de flujo Lazo, esto significa en un primer momento se ejecutara el subproceso descompuesto por el patrón lazo de la izquierda y en un segundo momento el descompuesto por el patrón Lazo de la derecha.

El primer subproceso se descompone a su vez en una actividad llamada *crear\_contrato+start* y una actividad invisible, la cual representa que existe ausencia de información en el proceso. La actividad *crear\_contrato+start* en el lazo se va a ejecutar tantas veces como aparezca en el registro de eventos, dicha actividad siempre va a aparecer como actividad de inicio de la secuencia del subproceso general.

El segundo subproceso se descompone en dos actividades, *crear\_contrato+complete* y una actividad invisible que al igual que en el subproceso anterior representa ausencia de información. La actividad *crear\_contrato+complete* puede aparecer tantas veces como se

registre en el registro de eventos y la misma es la que culmina la secuencia general iniciada por la actividad *crear\_contrato+start* del lazo izquierdo.

### 3.1.1. Identificación de métricas para el análisis de procesos aplicando la perspectiva temporal

Los nodos hojas del árbol de variantes representan las actividades que se ejecutan en el proceso, además cada nodo muestra las métricas para el análisis de procesos aplicando la perspectiva temporal, las cuales son duración total, media, máxima y mínima.

Las métrica en el modelo aparecen en abreviaturas, en la siguiente tabla se muestra como se identifican cada una en el árbol de variantes.

Tabla 1: Identificación de métricas

<i>Métricas</i>	<i>Identificador</i>
<i>Duración total</i>	T
<i>Duración media</i>	media
<i>Duración máxima</i>	Max
<i>Duración mínima</i>	Min

La métricas aparecen en el formato D HH: mm: ss., d representa la cantidad de días, h las horas, m los minutos y s los segundos de ejecución de cada actividad.

En la Tabla 2 se identifican las métricas para el modelo del proceso Procesar Solicitudes que se muestra en la figura 3.1 en la sección 3.1.

Tabla 2: Identificación de las métricas para el modelo del proceso Procesar Solicitudes

<b>Actividad</b>	<b>Código</b>	<b>Duración total</b>	<b>Duración media</b>	<b>Duración mínima</b>	<b>Duración máxima</b>

## CAPÍTULO 3: VALIDACIÓN DE LA SOLUCIÓN PROPUESTA

<b><i>crear_contrato+ complete</i></b>	d	7 minutos y 57 segundos	1 minutos y 59 segundos	2 minutos y 59 segundos	1 minutos y 17 segundos
<b><i>crear_cronograma_de _entrega+star</i></b>	e	1 minutos y 28 segundos	1 minutos y 28 segundos	1 minutos y 28 segundos	1 minutos y 28 segundos
<b><i>crear_contrato+star</i></b>	b	7 minutos y 57 segundos	1 minutos y 59 segundos	2 minutos y 59 segundos	1 minutos y 17 segundos
<b><i>crear_cronograma_de_entrega+ complete</i></b>	c	1 minutos y 28 segundos	1 minutos y 28 segundos	1 minutos y 28 segundos	1 minutos y 28 segundos

La tabla 3.2 se interpreta de la siguiente forma por cada una de las actividades. La actividad *crear\_contrato+ complete* codificada como d, tiene una duración total de 7 minutos y 57 segundos, una duración media de 1 minutos y 59 segundos, una duración máxima de 2 minutos y 59 segundos y duración mínima de 1 minutos y 17 segundos.

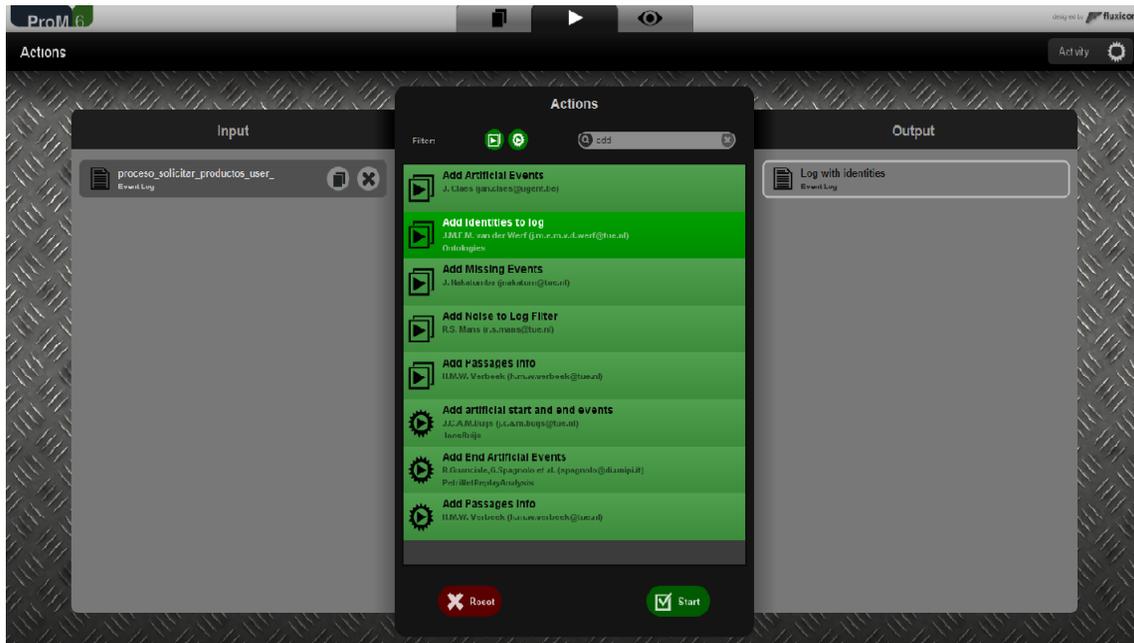
### 3.2. Validación de las actividades

El objetivo que se persigue realizando esta prueba es comprobar que el modelo generado por el componente muestra la cantidad de actividades existente en el registro de eventos de un proceso y que las actividades son las que realmente se ejecutan en el mismo.

Para la realización de esta prueba se utilizó el plugin *Add Identities to log* de ProM, pues permite mostrar información sobre los procesos de un registro de eventos, como por ejemplo muestra la cantidad de actividades que se ejecutan en un proceso, así como la ocurrencia absoluta y relativa de cada una de estas. Se empleó el mismo registro de evento utilizado en la sección 3.1, se seleccionó el proceso Solicitar Productos entre las fechas 2/05/2011 y 16/05/2013, como ejemplo para realizar la prueba.

A continuación se muestra el procedimiento realizado:

Se selecciona como entrada del plugin antes mencionado, el proceso Solicitar Productos como se muestra en la **Fig. 3.4**.



**Figura 3.4:** Plugin *Add Identities to log* de ProM

Fuente: Herramienta ProM

El plugin devuelve la información sobre el proceso Solicitar Productos como se muestra en la **Fig. 3.5**, donde se evidencia que el proceso, tienen un total de 13 actividades ejecutadas en el registro de eventos seleccionado.

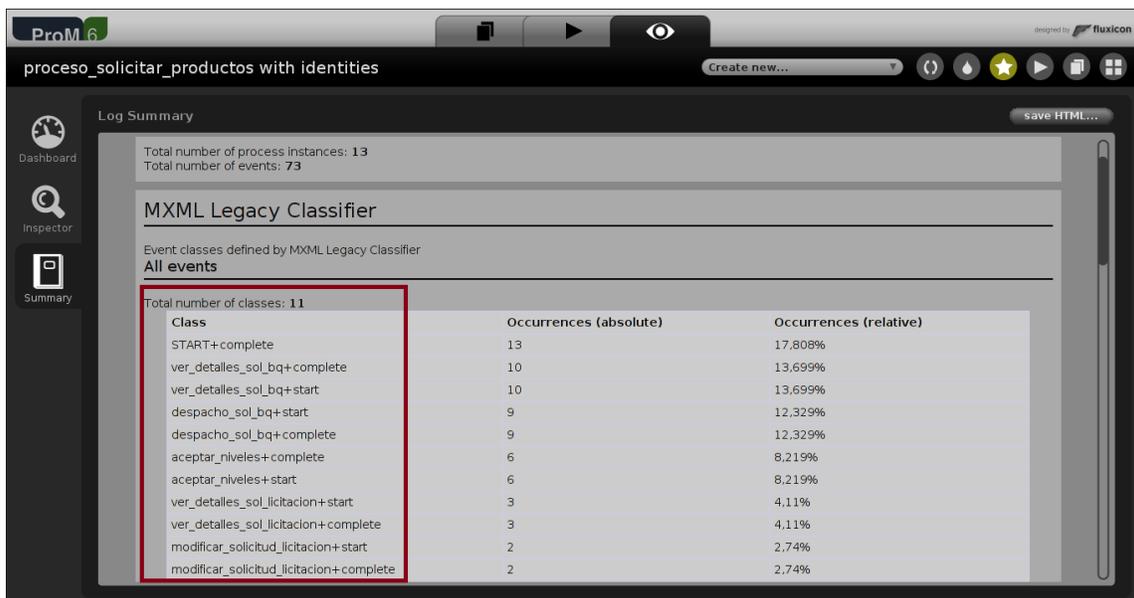
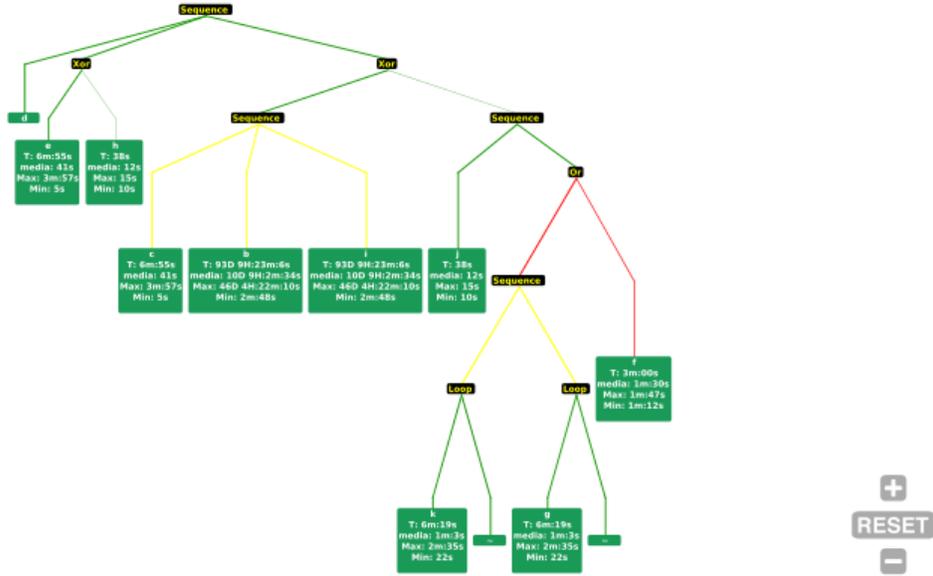


Figura 3.5: Cantidad de actividades del proceso Solicitar Productos

Fuente: Herramienta ProM

A continuación se muestra el modelo generado por el componente desarrollado desde el HIS.

Perspectiva Temporal



Leyenda de actividades		
Código	Actividad	Frecuencia
f	modificar_solicitud_licitacion+complete	2
g	aceptar_niveles+complete	6
d	START+complete (actividad de inicio de traza, sin valor para el modelo)	13
e	ver_detalle_sol_bq+start	10
b	despacho_sol_bq+start	9
c	ver_detalle_sol_bq+complete	10
a	modificar_solicitud_licitacion+start	2
j	ver_detalle_sol_licitacion+complete	3
k	aceptar_niveles+start	6
h	ver_detalle_sol_licitacion+start	3
i	despacho_sol_bq+complete	9

PATRONES

- SEQUENCE (Secuencia)
- XOR (Selección Exclusiva)
- OR (Selección No Exclusiva)
- PARALLELISM (Paralelismo)
- LOOP (Lazo)

Figura 3.6: Modelo del proceso Solicitar Productos

Fuente: Elaboración propia

La Fig. 3.6 evidencia como el modelo que se representa coincide con la cantidad de actividades devueltas por el plugin *Add Identities to log* de ProM y se puede apreciar que las actividades mostradas en la Leyenda de actividades del componente coinciden con las actividades mostradas por el plugin, demostrándose así la veracidad de la prueba realizada.

### **Conclusiones parciales**

El análisis de los resultados de la aplicación del componente desarrollado a un registro de eventos proveniente del módulo Almacén del HIS, permitió identificar en el modelo generado los patrones de control de flujo presentes en el registro de eventos y como se contextualizan a partir de la descomposición en subprocesos, además permitió identificar las métricas para el análisis de procesos aplicando la perspectiva temporal.

La validación de las actividades demostró que el componente muestra la cantidad de actividades existentes en el registro de eventos de un proceso y que las actividades mostradas son las que realmente se ejecutan en el mismo.

## CONCLUSIONES

Con la realización del presente trabajo de diploma se ha cumplido con el objetivo general propuesto, así como con las tareas de la investigación, arribándose a las siguientes conclusiones:

- El análisis realizado a las principales soluciones existentes tanto a nivel nacional como internacional, evidenció que las mismas presentan limitantes las cuales impiden su utilización en la presente investigación, por lo que los autores de la presente investigación desarrollaron un componente aplicando la perspectiva temporal de Minería procesos.
- El análisis de los principales algoritmos y técnicas de la Minería de procesos, evidenció que la técnica Minería de variantes, es una de las técnicas más completas del descubrimiento de proceso, pues permite la identificación de los patrones de control de flujo, el tratamiento y visualización del ruido y la ausencia de información.
- La asimilación de las herramientas y tecnologías propuestas, permitió obtener un componente de software para el análisis de procesos aplicando la perspectiva temporal.
- El desarrollo de un componente aplicando la perspectiva temporal de Minería de procesos permitió obtener las métricas necesarias para realizar el análisis de procesos aplicando la perspectiva temporal.
- La personalización del plugin *Variants Miner*, posibilitó obtener un árbol de variantes donde se visualiza en los nodos hojas las métricas para realizar el análisis de procesos aplicando la perspectiva temporal.
- La validación del componente desarrollado demostró que los resultados obtenidos tras su ejecución son correctos y que puede ser utilizado por analistas de procesos hospitalarios.

## RECOMENDACIONES

Al finalizar la presente investigación se recomienda:

- Mostrar la duración media, máxima, mínima entre las aristas del árbol de variantes que muestra el componente desarrollado.
- Incorporar una interfaz al componente, donde se explique el modelo generado por el mismo, en lenguaje natural, logrando este un mayor entendimiento sobre el funcionamiento y aumentando de esta manera la simpatía para su uso.

---

**REFERENCIAS BIBLIOGRÁFICAS**

- Pérez Alfonso, D., Yzquierdo Herrera, R., Silverio Castro, R., & Lazo Cortés, M. (2013). *Utilización de técnicas de minería de proceso en el entorno empresarial cubano*. Ciudad de la Habana, Cuba.
- The Apache Software Foundation*. (2012). Recuperado el 9 de 2 de 2014, de <http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/221/A6.pdf?sequence=>
- Alfaro, F. M. (2011). *Instituto Nacional de Estadística e Informática*. Recuperado el 2014 de 2 de 6, de <http://www.inei.gob.pe/biblioineipub/bancopub/Inf/Lib5103/Libro.pdf>
- Alfonso, D. P. (2014). *TÉCNICA PARA EL DIAGNÓSTICO DE VARIANTES DE PROCESOS DE NEGOCIO*. Ciudad de la Habana.
- Álvarez, A. L., Mora, M. M., Hernández, W. G., & Alemán, L. A. (19 de junio de 2013). *DESARROLLO DE LA ESPECIALIDAD PSICOLOGÍA DEL MÓDULO CONSULTA EXTERNA DEL SISTEMA*. Obtenido de <http://www.google.com/cu/url?sa=t&rct=j&q=JBoss+Seam++es+un+framework+que+integr+la+capa+de+presentaci%C3%B3n+%28JSF%29+con+la+capa+de+negocios+y+persistencia+%28JB%29%2C+funcionando%2C+seg%C3%BAn+versa+su+significado+en+espa%C3%B1ol%2C+como+una+%E2%80%A2>
- Aytulun, S. a. (2008). Business process modelling with stochastic networks. *International Journal of Production Research*, 46(10), 2743-2764.
- Chamorro Ahumada, M. C., & Maturana Valderrama, S. (s.f.). *Método para Aplicar Minería de procesos a la Distribución de Bebestibles No Alcohólicos*.
- Consejería de Hacienda y Administración Pública*. (2007). Recuperado el 9 de 2 de 2014, de <https://ws024.juntadeandalucia.es/ae/descargar/3191>.
- Dios Rubio, M. A., Framiñán Torres, J. M., Domínguez Cañizares, R., & León Blanco, J. M. (2010). Modelado y análisis de un proceso quirúrgico mediante técnicas de Minería de procesos. *Asociación para el Desarrollo de la Ingeniería de Organización - ADINGOR*, 243-257.
- Franky, C. .. (2010). *Java EE 5 (sucesor de J2EE)*. . Obtenido de [http://www.acis.org.co/fileadmin/Conferencias/ConfConsueloFranky\\_Abr19.pdf](http://www.acis.org.co/fileadmin/Conferencias/ConfConsueloFranky_Abr19.pdf)
- Gaguancela, V. D. (2012). *Universidad central del Ecuador*. Recuperado el 9 de 2 de 2014, de <http://www.dspace.uce.edu.ec/bitstream/25000/478/1/T-UCE-0011-27.pdf>.
- IEEE Task Force On Process Mining. (2011). *Manifiesto sobre Minería de Proceso*. Recuperado el 1 de 2 de 2015, de <http://www.win.tue.nl/ieeetfpm/lib/exe/fetch.php?media=shared:pmm>
- Ihaddadene. (2008). Extraction of business process models from workflow events logs. *International Journal of Parallel, Emergent & Distributed Systems*(23), 247-258.

- Industry., A. B. ( 2009). Pappa, D., N. Faltin, and V. Zimmermann. *International Journal of Advanced Corporate Learning*, 2, 16-24.
- (2006). *Introducción a JasperReports e iReport (Primera parte)*. .
- Introducción a JSF (Java Server Faces*. (2012). Obtenido de ]  
<http://www.desarrolloweb.com/articulos/2380.php>
- Ivar Jacobson, G. B. (2009). *El Proceso Unificado de Desarrollo de Software*. España. España.
- Jan Claes, I. V. (2012). *A visual analysis of the process of process modelig*.
- Jan Claes, I. V. (s.f.). *A visual analysis of the process of process modeling*.
- Leoni, M., W.M.P, & van der Aalst, . (2012). *Aligning Event Registros and Process Models for Multi- Perspective Conformance Checkingg: An Approach Based on Integer Linear Programming*. Netherlands: Eindhoven University of Technoregistry.
- Lucifer, P. (s.f.). *Java Runtime Environment - JRE*. Obtenido de  
<http://www.elleonplateadodeojosrojos.es/blog/java-runtime-environment-jre/>
- Manager, ARIS Process Performance. (2011). *ARIS Process Performance*. Obtenido de  
[http://www.softwareag.com/corporate/products/aris\\_platform/aris\\_controlling/aris\\_processes\\_performance/overview/default.asp](http://www.softwareag.com/corporate/products/aris_platform/aris_controlling/aris_processes_performance/overview/default.asp)
- Medeiros, A. (2006). *Genetic Process Mining*. Eindhoven University of Technoregistry.
- Oquendo, Y. Y. (2011). *Desarrollo de un componente de Transmisión de Audio y Video para el Sistema de Teleconsulta*. La Habana.
- Pérez Alfonso Damián, R. Y. (7 de Julio de 2014). *Utilización de técnicas de minería de proceso en el entorno empresarial cubano*. Recuperado el 8 de octubre de 2014, de  
[http://www.evolucion.cl/resumenes/Resumen\\_libro\\_Gesti%F3n\\_de\\_procesos\\_JBC\\_2011.pdf](http://www.evolucion.cl/resumenes/Resumen_libro_Gesti%F3n_de_procesos_JBC_2011.pdf)
- Process Mining Group, Math&CS department, Eindhoven University of Technology. (26 de 05 de 2009). *process mining*. Obtenido de  
[www.processmining.org\\_online\\_dottedchartanalysis.html](http://www.processmining.org_online_dottedchartanalysis.html)
- Ricardo Cruz, I., & Perdomo García, R. (2014). *Procedimiento de Conformidad aplicando Minería de procesos desde la perspectiva tiempo en el Sistema de Información Hospitalaria del Centro de Informática Médica*.
- Rogers, S. (2011). *Data is Scaling BI and Analytics-Data Growth is About to Accelerate Exponentially - Get Ready*. Information and Management. *nformation and Management - Brookfield*, 21(5), 14.
- Rondón, L. (3 de abril de 2009). *JAVA J2EE. JPA - Java Persistence API*. Obtenido de  
<http://www.luchorondon.blogspot.com/2009/04/jpa-java-persistence-api.html>.
- Rozinat, A. (2012 ). *Fluxicon. Process mining for professionals*. Obtenido de  
<http://fluxicon.com/disco/>
- Sánchez, J. M. (2008 ). *Migración de JSP a Facelets*. Obtenido de  
<http://www.adictosaltrabajo.com/tutoriales/tutoriales.php>

- Santiesteban, I. P., & Medina, M. R. (2011.). *Serie Científica Universidad de las Ciencias Informáticas*. Recuperado el 9 de 2 de 2014, de <http://publicaciones.uci.cu/index.php/SC>.
- Scribd. (2012). Obtenido de <http://es.scribd.com/doc/99836434/Persistencia-Hibernate>.
- van der Aalst, W. (2011). *Process Mining. Discovery, Conformance and Enhancement of Business Process*. London New York: Springer.
- Van der Aalst, W. M. (2009). *Conceptual Model for On Line Auditing. tech. report*. Obtenido de [BPMcenter.org](http://BPMcenter.org)
- Varga, J. L. (2007). *Biblioteca Universidad Complutense*. Recuperado el 14 de 2 de 2014, de <http://eprints.ucm.es/9064/1/TC2007-32.pdf>.
- Vergidis, K. C. (2008). Business process perspectives: Theoretical developments vs. real-world practice. *International Journal of Production Economics*, 114(1), 91-104.
- Weske, M. W. (2004). Advances in business process management. *Data & Knowledge Engineering*, 50(1), 1-8.
- Weske, M. W. (2004). Verbeek, Advances in business process management. *Data & Knowledge Engineering*, 50(1), 1-8.
- Wil van der Aalst, A. A. (2011). *Manifiesto sobre Minería de procesos*.
- Wojtkowski, W. e. (2009). Building a Process Performance Model for Business Activity Monitoring, in *Information Systems Development*. Springer US, 237-248.
- Yzquierdo Herrera, R. (2013). Minería de proceso como herramienta para la auditoría. *Ciencias de la Informació*, 44(2), 25-32.
- Zaninotto, F. y. (2008). *Symfony la Guía Definitiva*. New York: ISBN-13: 978-1590597866.

---

**BIBLIOGRAFÍA**

- Data is Scaling BI and Analytics-Data Growth is About to Accelerate Exponentially - Get Ready. Information and Management. Rogers, S. 2011. 5, 2011, Information and Management - Brookfield, Vol. 21, pág. 14.
- HERNÁNDEZ, PAOLA. 2012. APLICACIÓN DE TÉCNICAS DE MINERÍA DE PROCESOS (PM) PARA EL CONTROL Y MEJORAMIENTO DEL PROCESO DE COMPRAS NACIONALES E INTERNACIONALES DE BIENES Y SERVICIOS PARA PROYECTOS. BOGOTÁ, D.C. : s.n., 2012.
2012. The Apache Software Foundation. [En línea] 2012. [Citado el: 9 de 2 de 2014.] <http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/221/A6.pdf?sequence=> .
- Advances in business process management. Data & Knowledge Engineering. Weske , M., W.M.P., van der Aalst, y Verbee, H.M.W. 2004. 1, 2004, Vol. 50, págs. 1-8.
- Alfaro, Félix Murillo. 2011. Instituto Nacional de Estadística e Informática. [En línea] 2011. [Citado el: 2014 de 2 de 6.] <http://www.inei.gov.pe/biblioineipub/bancopub/Inf/Lib5103/Libro.pdf>.
- Alfonso, Damián Pérez. 2014. TÉCNICA PARA EL DIAGNÓSTICO DE VARIANTES DE PROCESOS DE NEGOCIO. Ciudad de la Habana : s.n., 2014.
- Almaguer, Norge . 2011. Documento normativo de Arquitectura de Software. 2011.
- Álvarez, Amaya Lorenzo, y otros. 2013. DESARROLLO DE LA ESPECIALIDAD PSICOLOGÍA DEL MÓDULO CONSULTA EXTERNA DEL SISTEMA . [En línea] 19 de junio de 2013. <http://www.google.com/cu/url?sa=t&rct=j&q=JBoss+Seam++es+un+framework+que+integra+la+capa+de+presentaci%C3%B3n+%28JSF%29+con+la+capa+de+negocios+y+persistencia+%28JB%29%2C+funcionamiento%2C+seguridad+aniversario+su+significado+en+espa%C3%B1ol%2C+como+una+...>
- autores, Colectivo de. 2005. Modelos para implantar la mejora continua en la gestión. 2005.
- Building a Process Performance Model for Business Activity Monitoring, in Information Systems Development. Wojtkowski, W., et al. 2009. 2009, Springer US, págs. 237-248.
- Business process modelling with stochastic networks. Aytulun, S.K. and A.F. Guneri. 2008. 10, 2008, International Journal of Production Research, Vol. 46, págs. 2743-2764.
- Business process perspectives: Theoretical developments vs. real-world practice. Vergidis, K., C.J. Turner, and A. Tiwari. 2008. 1, 2008, International Journal of Production Economics, Vol. 114, págs. 91-104.
- Chamorro , María Carolina y Valderr, Maturana . 2013. Método para Aplicar Minería de procesos a la Distribución de Bebestibles No Alcohólicos. 2013.

- 
2007. Consejería de Hacienda y Administración Pública. [En línea] 2007. [Citado el: 9 de 2 de 2014.] <https://ws024.juntadeandalucia.es/ae/descargar/3191..>
- Cruz, Ricardo Idalberto y Perdomo , Ramsés. 2014. Procedimiento de Conformidad aplicando Minería de procesos desde la perspectiva tiempo en el Sistema de Información Hospitalaria del Centro de Informática Médica. 2014.
- Extraction of business process models from workflow events logs. Ihaddadene. 2008. 23, 2008, International Journal of Parallel, Emergent & Distributed Systems, págs. 247-258.
- Franky, C. . 2010. Java EE 5 (sucesor de J2EE). . [En línea] 2010. [http://www.acis.org.co/fileadmin/Conferencias/ConfConsueloFranky\\_Abr19.pdf](http://www.acis.org.co/fileadmin/Conferencias/ConfConsueloFranky_Abr19.pdf).
- Gaguancela, Victor David Espinosa Vallejo y Alfonso Gustavo Gaguancela. 2012. Universidad central del Ecuador. [En línea] 2012. [Citado el: 9 de 2 de 2014.] <http://www.dspace.uce.edu.ec/bitstream/25000/478/1/T-UCE-0011-27.pdf..>
- Genetic Process Mining. Eindhoven University of Technoregistry. Medeiros, A.K. 2006. 2006.
- Gonzalez, Ing. Hugo. 2015. ENFOQUE BASADO EN PROCESOS COMO PRINCIPIO DE GESTIÓN. Calidad y Gestion. [En línea] 11 de 03 de 2015. <https://calidadgestion.wordpress.com/>.
- Hernández, Paola. 2012. Aplicación de técnicas de Minería de procesos (PM) para el control y mejoramiento del procesos de comras nacionales e internacionales de bienes y servicios para proyectos. BOGOTÁ, D.C. : s.n., 2012.
- IEEE Task Force On Process Mining. 2011. Manifiesto sobre Minería de Proceso. [En línea] 2011. [Citado el: 1 de 2 de 2015.] <http://www.win.tue.nl/ieeetfpm/lib/exe/fetch.php?media=shared:pmm>.
- Industry., Applying Business Process Oriented Learning. Pappa, D., Faltin, N. y Zimmermann, V. 2009. 2009, International Journal of Advanced Corporate Learning, Vol. 2, págs. 16-24.
2012. Introducción a JSF (Java Server Faces. [En línea] 2012. ] <http://www.desarrolloweb.com/articulos/2380.php>.
- Ivar Jacobson, Grady Booch y James Rumbaugh. 2009. El Proceso Unificado de Desarrollode Software. España. España : s.n., 2009.
- Jan Claes, Irene Vanderfeesten, Jakob Pinggera, Hajo A. Reijers, Barbara Weber and Geert Poels. 2012. A visual analysis of the process of process modelig. 2012.
- Jan Claes, Irene Vanderfeesten, Jakob Pinggera, Hajo A. Reijers, Barbara Weber and Geert Poels1. A visual analysis of the process of process modeling.
- Leoni, M, W.M.P y van der Aalst, . 2012. Aligning Event Registros and Process Models for Multi-Perspective Conformance Checking: An Approach Based on Integer Linear Programming. Netherlands : Eindhoven University of Technoregistry, 2012.
- LI C., REICHERT M., A WOMBACHER.,. 2011. Mining business process variants: Challenges, scenarios, algorithms. 2011. págs. 409-434. Vol. 70.

- Lucifer, P. Java Runtime Environment - JRE. [En línea] <http://www.elleonplateadodeojosrojos.es/blog/java-runtime-environment-jre/>.
- Manager, ARIS Process Performance. 2011. ARIS Process Performance. [En línea] 2011. [http://www.softwareag.com/corporate/products/aris\\_platform/aris\\_controlling/aris\\_processes\\_performance/overview/default.asp](http://www.softwareag.com/corporate/products/aris_platform/aris_controlling/aris_processes_performance/overview/default.asp).
- Minería de proceso como herramienta para la auditoría. Yzquierdo, Raykenler. 2013. 2, 2013, Ciencias de la Información, Vol. 44, págs. 25-32.
- Modelado y análisis de un proceso quirúrgico mediante técnicas de Minería de procesos. Dios, Manuel Alejandro; y otros. 2010. Sevilla : s.n., 2010, Asociación para el Desarrollo de la Ingeniería de Organización - ADINGOR, págs. 243-257.
- Manuel Alejandro Dios Rubio, José Manuel Framiñán Torres, Roberto Domínguez, José Miguel León Blanco. 2010. 2010, Asociación para el desarrollo de la ingeniería de organización, págs. pag. 243-257.
- Nariño, Arianys Hernández. 2010. Contribución a la gestión y mejora de. Matanzas : s.n., 2010.
- Oquendo, Yisel Ybarra Cristiá y Eddy Eliceo Alvarado. 2011. Desarrollo de un componente de Transmisión de Audio y Video para el Sistema de Teleconsulta. La Habana : s.n., 2011.
- Pérez Alfonso Damián, Raykenler Yzquierdo Herrera, Rogelio Silverio Castro<sup>1</sup>, Manuel Lazo Cortés. 2014. Utilización de técnicas de minería de proceso en el entorno empresarial cubano. [En línea] 7 de Julio de 2014. [Citado el: 8 de octubre de 2014.] [http://www.evolucion.cl/resumenes/Resumen\\_libro\\_Gesti%F3n\\_de\\_procesos\\_JBC\\_2011.pdf](http://www.evolucion.cl/resumenes/Resumen_libro_Gesti%F3n_de_procesos_JBC_2011.pdf).
- Pérez, Damián Alfonso, y otros. 2013. Utilización de técnicas de minería de proceso en el entorno empresarial cubano. Ciudad de la Habana, Cuba : s.n., 2013.
- Process Mining Group, Math&CS department, Eindhoven University of Technology. 2009. process mining. [En línea] 26 de 05 de 2009. [www.processmining.org\\_online\\_dottedchartanalysis.html](http://www.processmining.org_online_dottedchartanalysis.html).
- Process Mining. Discovery, Conformance and Enhancement of Business Process. van der Aalst, W.M.P. 2011. London New York: Springer : s.n., 2011.
- RAMÍREZ, PEDRO DANIEL HIDALGO. 2013. MODELO DE GESTIÓN Y ADMINISTRACIÓN DE PROYECTOS OPERACIONALES. SANTIAGO DE CHILE : s.n., 2013.
- Rondón, Luis. 2009. JAVA J2EE. JPA - Java Persistence API. [En línea] 3 de abril de 2009. <http://www.luchorondon.blogspot.com/2009/04/jpa-java-persistence-api.html>.
- Rozinat, Anne. 2012. Fluxicon. Process mining for professionals. [En línea] 2012. <http://fluxicon.com/disco/>.
- Sánchez, José Manuel Suárez. 2008. Migración de JSP a Facelets. [En línea] 2008. <http://www.adictosaltrabajo.com/tutoriales/tutoriales.php>.
- Santiesteban, Irina Ivis Pérez y Medina, Miguel Ramír. 2011.. Serie Científica Universidad de las Ciencias Informáticas. [En línea] 2011. [Citado el: 9 de 2 de 2014.] <http://publicaciones.uci.cu/index.php/SC..>

- 
- Scribd. 2012. [En línea] 2012. <http://es.scribd.com/doc/99836434/Persistencia-Hibernate..>
- SEPPE K.L.M, VANDEN BROUCKE, JAN VANTHIENEN, BART BAESENS. 2013. Third International Business Process Intelligence Challenge (BPIC'13): Volvo IT Belgium VINST. 2013.
- Thomas Molka, Wasif Gilani and Xiao-Jun Zeng. 2013. Dotted Chart and Control-Flow Analysis for a. Manchester : s.n., 2013.
- UPV, Gerencia. 2011 . SISTEMA DE GESTIÓN DE CALIDAD UPV. Valencia : s.n., 2011 .
- Van der Aalst, W. M. P., Hee, K. v., Werf, J. M. v. d., Kumar, A., & Verdonk, M. . 2009. Conceptual Model for On Line Auditing. tech. report. [En línea] 2009. BPMcenter.org.
- Varga, José Luis García Sarasa y Gloria Pérez. 2007. Biblioteca Universidad Complutense. [En línea] 2007. [Citado el: 14 de 2 de 2014.] <http://eprints.ucm.es/9064/1/TC2007-32.pdf..>
- Vázquez, Cristóbal . 2006. Introducción a JasperReports e iReport (Primera parte). 2006.
- Verbeek, Advances in business process management. Weske, M., van der Aalst, Van y H.M.W. 2004. 1, 2004, Data & Knowledge Engineering, Vol. 50, págs. 1-8.
- W.M.P. van de r Aalst, M.H. Schonenb erg.M. Song. 2012. Time Prediction Based on Pro cess Mining. Eindhoven : s.n., 2012.
- Wil van der Aalst, Arya Adriansyah, Ana Karla Alves de Medeiros, Franco Arcieri, Thomas. 2011. Manifiesto sobre Minería de procesos. 2011.
- Zaninotto, François y Potencier, Fabien. 2008. Symphony la Guía Definitiva. . New York : ISBN-13: 978-1590597866, 2008.

## ANEXOS

### Anexo 1. Especificación del requisito funcional: Generar modelo de proceso

Tabla 3: Especificación del requisito funcional: General modelos de proceso

Nº	Nombre	Descripción	Prioridad para el cliente	Complejidad	Referencias cruzadas
1	Generar modelo de proceso	Permite obtener un modelo de proceso a partir de configurar los parámetros requeridos y seleccionar el tipo de análisis.	Alta	Alta	Ver: CESIM_SIGEC_Especificacion_de_casos_de_uso_Generar modelo de proceso. CU 1

### Anexo 2. Especificación del caso de uso: Generar modelo de proceso

#### Diagrama de casos de uso



Figura anexo 2: Diagrama de casos de uso Generar modelo de proceso.

Fuente: Elaboración propia.

#### Especificación del caso de uso: Generar modelo de proceso

Tabla 4: CU 1. Generar modelo de proceso

<b>Objetivo</b>	Permite generar un modelo de proceso.
<b>Actores</b>	Analista de procesos hospitalarios.
<b>Resumen</b>	El caso de uso inicia cuando el actor accede a la opción Generar modelo de proceso, el sistema brinda la posibilidad de introducir los criterios para generar el modelo, el actor introduce los datos que considera como criterios para generar el modelo, el sistema a partir de los criterios seleccionados y consultando las entidades correspondientes genera el modelo de proceso. El caso de uso finaliza cuando queda generado el modelo de proceso.
<b>Complejidad</b>	Alta
<b>Prioridad</b>	Alta
<b>Precondiciones</b>	Que existan datos de los eventos almacenados en el sistema.
<b>Postcondiciones</b>	Queda generado el modelo de procesos.

**Flujo de eventos****Flujo básico**

	<b>Actor</b>	<b>Sistema</b>
	El caso de uso inicia cuando se accede a la opción Generar modelo de proceso	
		<p>Brinda la posibilidad de introducir los criterios para generar el modelo:</p> <ul style="list-style-type: none"> <li>• Procesos</li> <li>• Tipo de análisis</li> <li>• Desde</li> <li>• Hasta</li> </ul> <p>y permite:</p> <ul style="list-style-type: none"> <li>• Generar modelo de procesos</li> </ul> <p>Cancelar operación. Ver <b>Evento 1</b>: "Cancelar operación."</p>
	Introduce o selecciona los datos que considera como criterios para generar el modelo y	

	selecciona la opción de Generar modelos	
		<ul style="list-style-type: none"> <li>• Si el usuario deja valores obligatorios vacíos, ver <b>Evento 2:</b> "Existen campos obligatorios vacíos"</li> <li>• Si al buscar información para generar el modelo el sistema no encuentra ninguna, ver <b>Evento 3:</b> "No se encontró información."</li> </ul>
		<p>Genera el modelo a partir de los criterios seleccionados. Y permite:</p> <p>Ampliar el modelo. Ver <b>Evento 4:</b> "Ampliar modelo de proceso".</p>
		Termina el caso de uso

#### Flujos alternos

#### Evento 1. "Cancelar operación"

	Actor	Sistema
	Selecciona la opción de Cancelar operación.	
2.		Regresa a la vista anterior.
3.		Termina el caso de uso.

#### Evento 2. "Existen datos incompletos"

		Muestra un indicador sobre los campos incompletos.
		Regresa al paso 2 del <b>Flujo Normal de Eventos.</b>

#### Evento 3. "No se encontró información"

1.		Muestra el mensaje de información "No se encontró información que cumpla con los parámetros seleccionados."
2.		Regresa al paso 2 del <b>Flujo Normal de Eventos.</b>

---

**Evento 4: "Ampliar modelo de proceso"**

	Selecciona la opción de Ampliar	
		Amplía el modelo en otra ventana
1.		Termina el caso de uso
<b>Relaciones</b>	<b>CU incluidos</b>	No existen
	<b>CU extendidos</b>	

### Anexo 3: Diseño de caso de prueba del caso de uso: Generar modelo de procesos

Tabla 5: Escenario Generar modelo de procesos

Escenario	Descripción	Seleccionar procesos	Seleccionar tipo de análisis	Desde	Hasta	Respuesta del sistema	Flujo central
EC 1.1 Generar modelo de proceso	Se desea generar un modelo de procesos para analizar el mismo.	V	V	V	V	Muestra un modelo de proceso.	Muestra la interfaz para generar los modelos de procesos. Se introducen los datos. Se selecciona la opción Generar.
		Solicitar Productos	Perspectiva Temporal	10/11/2011	12/03/2015		
		V	V	V	V		
		Solicitar Productos	Incongruencias	08/10/2010	07/09/2012		
		V	V	V	V		
		Procesar Solicitudes	Ruido	12/05/2009	10/11/2011		
EC 1.2 Cancelar operación	Cancelar la opción de Generar modelo de proceso	NA	NA	NA	NA	Regresa a la vista anterior.	Muestra la interfaz para generar los modelos de procesos. Se introducen los datos. Se selecciona la opción

							Cancelar.
EC 1.2	Luego de haber introducido los datos, el sistema los verifica y valida, de haber datos incompletos, el sistema muestra un indicador sobre los campos incompletos.	V	/	/	V	Muestra un indicador sobre los campos incompletos.	Muestra la interfaz para generar los modelos de procesos. Se introducen los datos incompletos. Se selecciona la opción Generar. Muestra un indicador sobre los campos incompletos.
Existen		Procesar Solicitudes	Vacío	Vacío	12/03/2015		
datos		/	V	V	/		
incompletos		Vacío	Eventualidades	06/11/2012	Vacío		
		Procesar Solicitudes	Perspectiva Temporal	18/09/2015	13/02/2012		
		V	V	/	/		
		Solicitar Productos	Eventualidades	05/11/2011	02/02/2010		

Tabla 6: Descripción de las variables.

No	Nombre de campo	Clasificación	Valor Nulo	Descripción
1	Seleccionar procesos	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
2	Seleccionar tipo de análisis	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
3	Desde	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
4	Hasta	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.