



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Informática

Desarrollo de un módulo del estándar XES para el análisis de datos y estudio comparativo de distintas aplicaciones basadas en minería de procesos en un contexto clínico

Trabajo Fin de Grado

Grado en Ingeniería Informática

AUTOR/A: Bulnes Caicoya, Samuel José

Tutor/a: Ibáñez Sánchez, Gema

Cotutor/a: Bayo Montón, José Luis

CURSO ACADÉMICO: 2023/2024

Resum

Segons l'OMS, la població mundial cada vegada està més envellida, patint de més malalties cròniques, la qual cosa augmenta considerablement el gasto per pacient, i fa trontollar la sostenibilitat dels sistemes de salut.

Metodologies com Lean Six Sigma, poden contribuir a millorar la qualitat de vida de les persones a través de la millora contínua dels processos d'atenció, per exemple, reduint llistes d'espera o temps fins a l'atenció - a vegades tan crítics en malalties temps dependents com ictus -, contribuint també, a la sostenibilitat dels sistemes de salut.

Gran part d'esta tasca pot secundar-se en l'anàlisi de les dades disponibles en l'àmbit clínic. No obstant això, en la majoria dels casos, les dades dels quals es disposen no tenen la qualitat suficient com per a poder realitzar una anàlisi dels processos adequat. Estos errors poden ser en gran manera humans, generats, per exemple, en introduir informació o durant el disseny dels sistemes d'informació, ja que poden ser molt complexos per als clínics. Seguint amb l'exemple de ictus, els clínics solen prioritzar l'atenció del pacient abans que la introducció de dades administratives en el sistema, la qual cosa no reflectiria la realitat.

Referent a això, estàndard com XES poden ajudar a dissenyar sistemes d'informació, de manera que resulte més senzill registrar les activitats dels pacients o assegurar-se d'arreglar la informació necessària per a esta classe d'anàlisi.

D'altra banda, una de les tecnologies més prometedores per a assistir a l'optimització de processos, i a metodologies com Lean Six Sigma, és la mineria de processos, que engloba una sèrie de tècniques, ferramentes i mètodes per a monitorar i millorar els processos.

Per tant, es pot considerar que la combinació de l'estàndard XES, juntament amb la utilització d'aplicacions basades en mineria de processos, pot ajudar a preparar el camí per a poder realitzar una anàlisi òptima dels processos clínics. Així doncs, en este treball final de grau es proposa la implementació d'un mòdul de l'estàndard XES, i el disseny d'un estudi comparatiu entre diferents aplicacions basades en mineria de processos que suporten este estàndard.

Paraules clau: sostenibilitat; sistemes de salut; optimització de processos; qualitat de dades; estàndard XES; mineria de processos

Resumen

Según la OMS, la población mundial cada vez está más envejecida, sufriendo de más enfermedades crónicas, lo que aumenta considerablemente el gasto por paciente, y hace tambalear la sostenibilidad de los sistemas de salud.

Metodologías como Lean Six Sigma, pueden contribuir a mejorar la calidad de vida de las personas a través de la mejora continua de los procesos de atención, por ejemplo, reduciendo listas de espera o tiempos hasta la atención - a veces tan críticos en enfermedades tiempo dependientes como ictus -, contribuyendo también, a la sostenibilidad de los sistemas de salud.

Gran parte de esta tarea puede apoyarse en el análisis de los datos disponibles en el ámbito clínico. No obstante, en la mayoría de los casos, los datos de los que se disponen no tienen la calidad suficiente como para poder realizar un análisis de los procesos adecuado. Estos errores pueden ser en gran medida humanos, generados, por ejemplo, al introducir información o durante el diseño de los sistemas de información, ya que pueden ser muy complejos para los clínicos. Siguiendo con el ejemplo de ictus, los clínicos suelen priorizar la atención del paciente antes que la introducción de datos administrativos en el sistema, lo que no reflejaría la realidad.

A este respecto, estándares como XES pueden ayudar a diseñar sistemas de información, de manera que resulte más sencillo registrar las actividades de los pacientes o asegurarse de recoger la información necesaria para esta clase de análisis.

Por otro lado, una de las tecnologías más prometedoras para asistir a la optimización de procesos, y a metodologías como Lean Six Sigma, es la minería de procesos, que engloba una serie de técnicas, herramientas y métodos para monitorear y mejorar los procesos.

Por tanto, se puede considerar que la combinación del estándar XES, junto con la utilización de aplicaciones basadas en minería de procesos, puede ayudar a preparar el camino para poder realizar un análisis óptimo de los procesos clínicos. Así pues, en este trabajo final de grado se propone la implementación de un módulo del estándar XES, y el diseño de un estudio comparativo entre distintas aplicaciones basadas en minería de procesos que soporten dicho estándar.

Palabras clave: sostenibilidad; sistemas de salud; optimización de procesos; calidad de datos; estándar XES; minería de procesos

Abstract

According to the WHO, the world's population is getting older, suffering from more chronic diseases, which considerably increases the cost per patient, and is jeopardizing the sustainability of healthcare systems.

Methodologies such as Lean Six Sigma can contribute to improving the quality of life of people through the continuous improvement of care processes, for example, reducing waiting lists or times to care - sometimes so critical in time-dependent diseases such as stroke -, also contributing to the sustainability of health systems.

Much of this task can be supported by the analysis of data available in the clinical setting. However, in most cases, the data available are not of sufficient quality to allow adequate process analysis. These errors can be largely human, generated, for example, when entering information or during the design of information systems, as they can be very complex for clinicians. Continuing with the stroke example, clinicians often prioritize patient care over entering administrative data into the system, which would not reflect reality.

In this respect, standards such as XES can help to design information systems so that it is easier to record patient activities or to ensure that the necessary information is collected for this kind of analysis.

On the other hand, one of the most promising technologies to assist process optimization, and methodologies such as Lean Six Sigma, is process mining, which encompasses a number of techniques, tools and methods for monitoring and improving processes.

Therefore, it can be considered that the combination of the XES standard, together with the use of applications based on process mining, can help pave the way to be able to perform an optimal analysis of clinical processes. So, this final degree work proposes the implementation of a module of the XES standard, and the design of a comparative study between different applications based on process mining that support this standard.

Key words: sustainability; healthcare systems; process optimization; data quality; XES standard; process mining

Índice general

Índice general	VII
Índice de figuras	IX
<hr/>	
1 Introducción	1
1.1 Motivación	2
1.2 Objetivos	2
1.3 Estructura de la memoria	3
2 Materiales y métodos	5
2.1 Metodologías para la mejora continua de procesos	5
2.1.1 Justificación elección estándar	7
2.2 Estándar XES	7
2.2.1 Descripción	7
2.3 Aplicaciones usadas para la comparativa	11
3 Resultados	19
3.1 Desarrollo módulos del estándar XES	19
3.2 Comparativa entre aplicaciones basadas en minería de procesos . .	20
3.2.1 Criterios de selección	20
3.2.2 Metodología usada para comparación	24
3.2.3 Resultados de la comparativa	24
3.2.4 Análisis de resultados	33
4 Relación con la carrera	53
5 Conclusiones y Trabajo Futuro	55
Bibliografía	57
<hr/>	
Apéndice	
A ODS	59

Índice de figuras

2.1	Diagrama de clases UML 2.0 para el metamodelo completo del estándar XES. Imagen proveniente del estándar XES [26]	15
2.2	Diagrama de flujo de la máquina de estados para el estándar XES. Imagen proveniente del estándar XES [26]	16
2.3	Fases previas al análisis de datos	17
2.4	Logo aplicación minería de procesos PMAApp	17
2.5	Logo aplicación minería de procesos Celonis	17
2.6	Logo aplicación minería de procesos ProM	17
2.7	Logo aplicación minería de procesos Disco	18
2.8	Logo aplicación minería de procesos Apromore	18
3.1	Estructura de clases	35
3.2	Parte del código desarrollado donde destaca la cabecera	36
3.3	Clase usada para el desarrollo del módulo Reader	36
3.4	Visualización módulo Reader en PMAApp	36
3.5	Clase utilizada para desarrollar saver	37
3.6	Otra clase utilizada para desarrollar el módulo saver	37
3.7	Visualización módulo Reader en PMAApp	38
3.8	Resultados Cuestiones Generales Comparativa	38
3.9	Resultados Cuestiones Manejo de calidad de Datos Comparativa	38
3.10	Resultados Cuestiones Técnicas Comparativa	38
3.11	Resultados Cuestiones Análisis Comparativa	39
3.12	Resultados Cuestiones Seguridad Comparativa	39
3.13	PMAApp - Pantalla inicial una vez se inicia la aplicación	39
3.14	PMAApp - Pantalla previa a la ejecución del análisis del archivo	40
3.15	PMAApp - Pantalla con un lector de archivos xes en la fase 1 y un tipo de algoritmo de descubrimiento del proceso en la fase 3	40
3.16	PMAApp - Pantalla con la ruta del archivo xes a leer, al algoritmo de descubrimiento del proceso, y otro algoritmo en la fase 4	41
3.17	PMAApp - Pantalla donde figura el proceso descubierto	41
3.18	PMAApp - Pantalla donde figura información de las trazas del proceso descubierto	42
3.19	PMAApp - Pantalla donde figura información relativa a cuestiones que se pueden modificar con respecto al realce del proceso obtenido	42
3.20	PMAApp - Pantalla donde figura información relativa a asuntos que se pueden modificar o agregar al proceso obtenido con respecto a estadísticas y gráficos	43
3.21	Celonis - Pantalla donde se carga el archivo xes para su análisis, siendo xes una extensión que acepta esta app	43

3.22 Celonis - Pantalla donde figura diversa información relativa al proceso	44
3.23 Celonis - Pantalla donde figura más información relativa al proceso	44
3.24 Celonis - Pantalla donde figuran las distintas etapas del análisis del archivo	44
3.25 Celonis - Preguntas disponibles para el usuario	45
3.26 Celonis - Información relativa al proceso analizado	45
3.27 Celonis - Información adicional del proceso analizado	45
3.28 Celonis - representación gráfica del proceso	46
3.29 Prom - Pantalla inicial de la aplicación donde se cargaría el archivo para su posterior análisis	46
3.30 Prom - Distintas opciones de análisis del archivo cargado	47
3.31 Prom - Salida gráficas del proceso analizado	47
3.32 Prom - Otra posible salida gráfica del proceso analizado	48
3.33 Disco - Imagen donde figura el proceso representado gráficamente	48
3.34 Disco - Imagen donde figura otro posible resultado a obtener como salida del análisis del proceso	49
3.35 Disco - Captura con datos estadísticos del proceso representado	49
3.36 Apromore - Pantalla inicial donde se carga el archivo para su posterior análisis	50
3.37 Apromore - Estadísticas del proceso obtenido como resultado	50
3.38 Apromore - Más estadísticas del proceso obtenido como resultado	50
3.39 Apromore - Proceso representado gráficamente	51
3.40 Apromore - Otra manera de representar el proceso gráficamente	51

CAPÍTULO 1

Introducción

Según la OMS, la población mundial cada vez está más envejecida, sufriendo de más enfermedades crónicas [1], lo que aumenta considerablemente el gasto por paciente, y hace tambalear la sostenibilidad de los sistemas de salud[1]. Esto se puede ver reflejado en el servicio de urgencias, por ejemplo, en el tiempo que cada una de las personas tiene que esperar desde que entra al servicio de urgencias del hospital hasta que es atendido, tiempo que cada vez es mayor debido al incremento de pacientes, aumentando así la estancia media de los mismos [16].

Aunque ya existen distintos estándares para priorizar a los pacientes según su tipo de urgencia, como la asignación del nivel de urgencia, cabe destacar que lo que falla realmente es la eficiencia de los procesos, ya que, por ejemplo, pongamos el caso hipotético de que un paciente con un nivel 5 de urgencia, nivel menos urgente, seguro que si que es triado correctamente, aunque si el proceso en sí no esté bien diseñado, pueden generarse tiempos de espera que no corresponden con una buena atención.

Hoy en día existen metodologías como Lean Six Sigma [17] que pueden ayudar a mejorar la calidad de atención, contribuyendo así, por ejemplo, a disminuir los tiempos de espera de los pacientes según sus niveles asignados, y en general, todo lo relacionado en el ámbito sanitario a los procesos de atención al cliente, que en este caso, son los pacientes que acuden a cada uno de los centros sanitarios existentes[2]. Estas mejoras son fundamentales para, por ejemplo, disminuir los tiempos de espera de aquellos pacientes que acuden con un nivel de urgencia máximo, que pueden poseer algún tipo de enfermedad o daño que debe ser tratado cuanto antes, ya que puede ser de vida o muerte, como puede ser la enfermedad denominada ictus -, enfermedad neurológica grave, pudiendo ocasionar graves lesiones al paciente que la padezca [3, 4, 14].

Gran parte de esto se puede apoyar en el análisis de los datos clínicos de los pacientes, donde en la mayoría de los casos, los datos disponibles no tienen la calidad necesaria para poder analizar cada uno de los procesos, y donde en gran parte, esta falta de calidad en los datos es debido a errores humanos durante la introducción de los datos o las fases de diseño de los sistemas de información. Por ejemplo, hablando de ictus, es fundamental conocer cuando se le administró al paciente que padece esta enfermedad el tratamiento, debido a que las dos primeras horas son las más importantes de cara a la recuperación del paciente,

pero en muchos casos esta información es introducida a posteriori dado que se prioriza la atención del paciente [5]

Teniendo en cuenta lo anteriormente comentado, también cabe destacar que el análisis de los datos puede ayudar a identificar puntos de mejora en estos mencionados sistemas de salud, pero, no obstante, la calidad de los mismos no suele ser buena, debido, principalmente, a errores humanos. En general y, para cerrar este punto, el objetivo principal es mejorar los procesos, utilizando para ello minería de procesos como tecnología para analizar los mismos. Por lo tanto, proponemos realizar una implementación de XES, que es un estándar para trabajar en minería de datos.

1.1 Motivación

Hoy en día, el análisis de datos clínicos no está exento de problemas relativos a la calidad de los mismos, lo que puede generar inconvenientes a la hora de analizarlos, teniendo como consecuencia, por ejemplo, toma de decisiones erróneas, los pacientes viéndose afectados negativamente durante su tratamiento [25].

Existen estándares que pueden contribuir a mejorar la calidad de los datos, como es el estándar XES [7], junto con el uso de aplicaciones basadas en minería de procesos [6], con el objetivo de conseguir que la calidad de la información ofrecida de cada uno de los pacientes sea la máxima posible, y que nos permita analizar uno a uno los procesos clínicos por los que pasan los pacientes y ver que sucede en cada uno de ellos, pudiendo identificar posibles cuellos de botella o cuestiones que se podrían optimizar.

El uso de este estándar, junto con las aplicaciones basadas en minería de procesos, mejoraría en gran medida la calidad de los datos ofrecidos de cada paciente y disminuiría los problemas ocasionados a los mismos, aumentando así la satisfacción del paciente con los servicios sanitarios ofrecidos y disminuyendo el número de veces que cada paciente visita a los servicios sanitarios anualmente.

La adopción de este estándar contribuiría a la mejora de la calidad de los datos, junto con el uso de aplicaciones basadas en minería de procesos, influiría principalmente en el ámbito sanitario relativo al análisis de datos clínicos y optimización de los mismos, como también influiría en lo relativo a la optimización de procesos.

1.2 Objetivos

El principal objetivo del desarrollo de este trabajo es el desarrollo de un módulo del estándar XES para la mejora de la calidad de los datos en ámbitos sanitarios.

Para poder llevar a cabo este desarrollo de un módulo del estándar XES, es importante destacar ciertos objetivos secundarios, como los siguientes:

1. **O1: El desarrollo de un módulo para la lectura de archivos que cumplan con el estándar XES.** Esto ayuda a la simplificación del procesado y análisis de los datos.
2. **O2: El desarrollo de un módulo para el guardado de archivos con extensión .xes.** Una vez procesados estos datos, se pueden exportar como xes, pudiendo así eliminar errores en los datos y pudiendo utilizar los mismos en otras aplicaciones que soporten xes.
3. **O3: El desarrollo de una comparativa entre las aplicaciones basadas en minería de procesos,** en donde los puntos claves a comparar son los siguientes (Cuestiones generales, cuestiones relativas a el manejo de la calidad de los datos, aspectos relativos a las técnicas, al análisis y a la seguridad). Gracias a esta comparativa, la cual se detalla más adelante en este trabajo, y entre los beneficios que pueden aportar a XES estas aplicaciones basadas en minería de procesos, se encontraría el poder dar una visión de cada una de ellas, y así elegir la que más se adapte a lo que se intenta conseguir, lo que contribuiría a que se consiga el objetivo principal.

1.3 Estructura de la memoria

En la memoria de este trabajo se tratarán varios aspectos acerca del estándar XES, el cual se usó para poder desarrollar este trabajo, como también asuntos acerca de la comparativa realizada. Entre los puntos que se van a tratar, se pueden mencionar los siguientes:

1. Un primer capítulo **1**, donde se comentan cuestiones relativas a la situación presente hoy en día con respecto a la calidad de los datos que se ofrecen y en qué afecta esto en el mundo sanitario, y el cómo el uso del estándar XES, junto con aplicaciones basadas en minería de procesos, puede mejorar esta situación.
2. Un segundo capítulo **2**, en donde se comenta qué elementos se han utilizado para llevar a cabo este trabajo, en este caso, acerca del estándar XES, mencionando qué es dicho estándar y que compone al mismo, además de una breve justificación del por qué se seleccionó el estándar XES como primera opción. Además, también se mencionan cuestiones acerca de la minería de procesos y de la metodología Lean Six Sigma.
3. Un tercer capítulo **3**, donde se comenta todo lo obtenido con el desarrollo de este trabajo, incluyendo principalmente cuestiones relativas a una comparativa entre aplicaciones basadas en minería de procesos, como los criterios de selección usados, la metodología usada para la comparativa, los resultados de la misma y un análisis de resultados final. Asimismo, también se resalta el cómo se llevó a cabo la implementación de este proyecto.
4. Un cuarto capítulo **4**, donde se destaca la relación de este trabajo con la carrera cursada, destacando todas aquellas tecnologías, entornos y lenguajes de programación que se vieron a lo largo de la misma.

5. Un último capítulo 5, donde se da un cierre general a todo el trabajo realizado, destacando las cuestiones más importantes.

CAPÍTULO 2

Materiales y métodos

2.1 Metodologías para la mejora continua de procesos

En el capítulo anterior hemos hablado de la necesidad de encontrar soluciones que contribuyan a la sostenibilidad de los sistemas de salud a través de la mejora de los procesos asistenciales. Metodologías como Lean Six Sigma buscan mejorar los procesos en las empresas, al combinar los principios de Lean Manufacturing y Six Sigma, con el objetivo de incrementar la rentabilidad y la productividad. Esta metodología se enfoca en eliminar los desperdicios y reducir la variabilidad en los procesos, asegurando que los productos y servicios se ajusten a los requisitos del cliente. Aunque esta metodología, inicialmente, surgió en el sector industrial, ha sido adoptado ampliamente en el sector de servicios y en empresas de todo el mundo que buscan mejorar su liderazgo en el mercado [2].

1. *Lean manufacturing*: Es un sistema de organización del trabajo que pone el foco en la mejora del sistema de producción. Para esto se basa en la eliminación de aquellas actividades que no aportan valor al proceso ni al cliente. Estas se denominan despilfarros o desperdicios, y son aquellas tareas que implican la sobreproducción, altos tiempos de espera o desperfectos en los productos, por ejemplo [20, 21].
2. *Six sigma*: Su objetivo es incrementar la calidad al minimizar la variabilidad en los procedimientos a través de técnicas estadísticas. Se fundamenta en la premisa de que los procedimientos deben satisfacer las exigencias del cliente, tratando cualquier discrepancia como un error que necesita corregirse [2].

Este procedimiento combina dos principios ya comentados, y lo hace mediante un método denominado DMAMC, que consiste en lo siguiente [2]:

1. *Definir*: Se identifica el problema, el valor para el cliente, y los objetivos del proyecto. También se establece el equipo de trabajo y se valida el problema a resolver.

2. *Medir*: Se mapea el proceso actual y se recopilan datos para determinar el rendimiento y fiabilidad, y para identificar aquellas áreas de mejora que puedan existir.
3. *Analizar*: Se identifica las causas raíz de la variación y los problemas en el proceso. Se centra en entender los factores que influyen en los resultados
4. *Mejorar*: Se implementan cambios en el proceso para optimizar el mismo.
5. *Controlar*: Se establecen controles para poder mantener y consolidar las mejoras implementadas.

Así pues, el concepto Six Sigma permite conocer y comprender los procesos, de tal forma que se puedan eliminar los desperdicios generados en ellos, siendo algunos de éstos, por ejemplo, defectos, esperas, sobreproducción, entre otros.

Por otro lado, el procedimiento comentado anteriormente, en combinación con la minería de procesos, puede significar una mejora en la optimización de procesos, ya que ambas se centran en el análisis de procesos y en la mejora de los mismos, por lo que a continuación se destaca lo más importante de lo relativo a la minería de procesos, antes de entrar en detalle en el capítulo, siendo esta misma un enfoque que analiza los datos desde los sistemas de TI para obtener información objetiva y descubrir problemas ocultos. Este enfoque se realiza mediante un software que combina ciencia de datos y gestión de procesos para descubrir y verificar la conformidad, además de mejorar los procesos reales [6].

Además de esto, este tipo de análisis de procesos utiliza análisis de datos para descubrir, supervisar y mejorar los procesos de negocios fundamentales, lo que funciona de la siguiente manera y lo que permite identificar semejanzas entre el objetivo de Lean Six Sigma y la minería de procesos, proporcionando soporte esta última a la metodología mencionada [6]:

1. *Recolección de datos*: El software de minería de procesos recolecta datos de varias fuentes, como registros de procesos, sistemas ERP y bases de datos. Estos datos contienen información acerca de los pasos, eventos y actores relacionados con los procesos de análisis
2. *Descubrimiento de procesos*: Los datos recolectados se utilizan para descubrir de manera automática los procesos subyacentes y crear una representación visual de ellos en un mapa de procesos.
3. *Análisis de procesos*: Los mapas de procesos son analizados para identificar ineficiencias y otros problemas de los procesos.
4. *Optimización*: Se implementan las mejoras necesarias al proceso una vez identificadas las ineficiencias.
5. *Monitorización*: Este software de la minería de procesos también puede servir para la vigilancia continua de los procesos, pudiendo detectar, por ejemplo, problemas con antelación, ayudando así a evitar que los mismos se puedan agravar.

Con esto ya mencionado, se puede notar claramente como tanto la metodología como la minería de procesos tienen un objetivo común, el cuál es la mejora de los procesos. Además de lo ya tratado, en lo relativo a la calidad de los datos, que en muchos casos no es óptima, se propone la utilización de XES como estándar para la gestión de datos clínicos, estándar pensado para reducir problemas en los datos y poder ser utilizados los mismos por aplicaciones de minería de procesos. Dicho estándar se detalla a partir de la siguiente subsección, empezando por el porqué de la selección del mismo para el desarrollo de este trabajo.

2.1.1. Justificación elección estándar

Después de una breve introducción a este capítulo, a continuación se va a comentar el porqué de la selección de este estándar para el desarrollo de este trabajo, ya que existían otras opciones, pero la que se consideró más adecuada fue la de este estándar, selección que se justificará a continuación.

En este caso, el estándar XES fue el seleccionado por los siguientes puntos principalmente [7]:

1. *Interoperabilidad y estandarización*: XES es un estándar bien reconocido y establecido en el ámbito de la minería de procesos. Al utilizarlo, garantiza la compatibilidad con numerosas herramientas y sistemas que también cumplen con este estándar.
2. *Tolerancia y extensibilidad*: XES permite a los usuarios personalizar el estándar según los requisitos específicos, y al mismo tiempo, poder compartir y analizar datos de manera consistente.
3. *Afinidad*: XES es ampliamente aceptado como el estándar para la minería de procesos, lo que lo convierte en la mejor opción para proyectos centrados en analizar y mejorar los procesos comerciales basados en eventos.
4. *Integración y análisis de datos*: XES permite la consolidación de eventos de diferentes fuentes en un único formato de registro de eventos, así como un análisis en profundidad crucial para la mejora.

2.2 Estándar XES

A continuación, se detalla el estándar XES, desde su composición, hasta aquellas cuestiones relevantes como que es obligatorio, que no, y la flexibilidad que posee este estándar, permitiendo, por ejemplo, agregar extensiones con el objetivo de agregar atributos adicionales. Además, también se comentarán cuestiones relacionadas a la aplicación principal utilizada para llevar a cabo este trabajo [7].

2.2.1. Descripción

El estándar XES establece una lingüística para un lenguaje basado en etiquetas, cuyo objetivo principal es el de proporcionar a los diseñadores de sistemas

de información una metodología unificada y ampliable para capturar los comportamientos de los sistemas, mediante registros y flujos de eventos. Este estándar incluye un .esquema XML" que describe la estructura de un registro/transmisión de eventos XES y un .esquema XML" que describe la estructura de una extensión de ese registro/transmisión. Además, este estándar incluye una selección fundamental de los prototipos de .extensión XES", que proporcionan semántica a ciertos rasgos registrados en el flujo de eventos o el registro de eventos [7].

Teniendo en cuenta la imagen correspondiente a la jerarquía XES (Figura 2.1), es importante mencionar que este estándar está compuesto principalmente por los siguientes elementos, los cuáles son [7]:

1. LOG: Componente que representa toda la información relacionada a cierto proceso, y que puede o no contener una colección de trazas, componente que se explica posteriormente, seguido de una lista de eventos, elemento que también se explica posteriormente, la cuál puede ser vacía, es decir, que no haya eventos presentes. También es importante mencionar que el orden de los eventos en la lista de eventos es importante, ya que es el orden en que fueron ocurriendo los mismos.
2. TRAZA: Componente que representa la ejecución de un solo caso de cierto proceso, y que puede contener o no una lista/colección de eventos relacionados con el caso que se está tratando del proceso. En este caso, al igual que antes, el orden de los eventos en la lista de los mismos es importante
3. EVENTO: Componente que representa un gránulo atómico de actividad que fue observado. Es evidente a que caso pertenece el evento si el mismo ocurre en alguna traza.
4. ATRIBUTO: Componente que proporciona información de los distintos componentes mencionados anteriormente, entre los que se pueden mencionar los siguientes tipos:
 - a) *Atributos elementales*: Son atributos que contienen un valor básico/único, como los atributos de tipo string, date, time, atributos de números enteros, atributos de números reales, atributos booleanos y atributos de tipo ID.
 - b) *Atributos compuestos*: Son atributos que pueden o no contener múltiples valores, como lo son las listas.
 - c) *Atributos globales*: Un log puede o no tener una lista global de declaraciones de este tipo de atributos, los cuáles son una característica requerida para el cumplimiento de este estándar, y pueden servir tanto para los rastros como para los eventos. Existen dos tipos de estos atributos, los cuáles son:
 - 1) *Atributos de evento globales*: Son atributos que están disponibles y definidos para cada evento en el registro, obteniendo como resultado que cada evento en el registro contenga un atributo con la clave y el tipo de datos dado, pero con posibilidad de que contenga

un valor válido diferente. El valor proporcionado para una declaración de este tipo de atributo solo es característico en caso de que sea necesario crear un evento para el cual no se proporciona valor alguno para ese atributo. En otro caso, el valor de la declaración es insignificante, ya que no se utilizará.

- 2) *Atributos de traza globales*: Son atributos que están disponibles y definidos para cada rastro en el registro, obteniendo como resultado que cada traza en el registro contenga un atributo con la clave y el tipo de datos dado, aunque con la posibilidad de que posea un valor distinto. El valor proporcionado para una declaración de este tipo de atributo solo es característico en caso de que sea necesario crear un rastro para el cual no se proporciona valor alguno para ese atributo. En otro caso, el valor de la declaración es insignificante, ya que no se utilizará.

5. **CLASIFICADORES**: Componente que asigna a cada evento una identidad, lo que permite que los mismos se puedan comparar entre sí en base a la identidad de cada uno. Algunos ejemplos de estas identidades pueden ser el nombre descriptivo del evento, el nombre descriptivo del caso al que el evento está relacionado, el nombre descriptivo de la causa del evento o el nombre descriptivo del caso relacionado al evento. Estos clasificadores pueden ser de dos tipos:

- a) *Clasificadores de evento*: Se define mediante una lista ordenada de claves de atributos. La identidad del evento se derivará de los valores reales de los atributos con estas claves. Un atributo cuya clave aparece en una lista de clasificadores de eventos se declarará como un atributo de evento global antes de que se defina el clasificador de eventos, ya que el clasificador de eventos requiere el valor real del atributo.
- b) *Clasificadores de traza*: Se define mediante una lista ordenada de claves de atributos. La identidad de cada traza se derivará de los valores reales de los atributos con estas claves. Un atributo cuya clave aparece en una lista de clasificadores de seguimiento se declarará como atributo de seguimiento global antes de que se defina el clasificador de seguimiento, ya que el clasificador de seguimiento requiere el valor real del atributo.

6. **EXTENSIONES**: Esta Norma no define un conjunto específico de atributos por componente. Como tal, la semántica de los atributos de datos que contienen estos elementos debe ser necesariamente ambigua, lo que dificulta la interpretación de esos datos. Esta ambigüedad se resuelve con el concepto de las extensiones en esta Norma. Una extensión define, para cada tipo de componente, un conjunto de atributos (posiblemente vacío). La extensión proporciona puntos de referencia para interpretar estos atributos y, por tanto, sus componentes. Por lo tanto, las extensiones son principalmente un vehículo para adjuntar semántica a un conjunto de atributos definidos por componente. Una extensión tendrá un nombre descriptivo, un prefijo y un URI. El prefijo es el prefijo de todos los atributos definidos por la extensión,

lo que significa que las claves de todos los atributos definidos por la extensión estarán precedidas por este prefijo y un carácter de separación de dos puntos (como un espacio de nombres en XML). El URI es un URI único que apunta a la definición de la extensión.

Con todo lo comentado anteriormente, en la imagen (Figura 2.2), se puede visualizar todo lo relacionado a los estados por los que pasa el estándar XES [7]:

Además de lo comentado acerca de la minería de procesos, también es importante tratar con distintos aspectos relativos a las aplicaciones basadas en minería de procesos que existen hoy en día en el mercado, y de las cuales se elegirá una para la implementación del estándar, aspectos que se comentan a continuación.

Actualmente, existen muchas apps basadas en minería de procesos, pero debido a la disponibilidad de las mismas, las más interesantes han sido: PMApp, Celonis, ProM, Disco y Apromore. Asimismo, la gran mayoría de estas aplicaciones están desarrolladas de manera general para que pueda ser utilizada en distintos ámbitos, por ejemplo, industria o banca. Esto es así a excepción de PMApp, que está diseñada especialmente para salud. Así pues, y dadas estas premisas, para la implementación del estándar XES se eligió PMApp, aplicación desarrollada en Visual Studio .NET, que sigue un diseño modular para facilitar la incorporación de nuevos recursos a través de plugins o librerías de una manera sencilla.

Para realizar el análisis de los datos, PMApp implementa cinco fases, las cuales permiten añadir recursos en forma de bloques *drag & drop*. Estos bloques pueden ser de 5 tipos diferentes (Figura 2.3) que a continuación se explican:

1. *Factorías*: Esta primera fase consiste en crear logs utilizando bloques de factoría. La figura 2.3 es un claro ejemplo de datos sintéticos de un departamento de emergencias, donde se puede identificar un bloque CSV ingestor de trazas, el cuál gestiona las fuentes de datos y establece las reglas para modelar el proceso en fases posteriores, siendo en este ejemplo leyendo un archivo CSV. [12].
2. *Filtering*: En esta segunda fase, luego de que el log se haya creado, se pueden aplicar filtros para corregir, dividir y agrupar trazas, según sea la necesidad en un momento específico, permitiendo así modificaciones antes del proceso de descubrimiento del proceso. Continuando con la figura 2.3, bloques como los indicadores del modelo LoSFilter (duración de la estancia) disponibles en lo relativo a ámbitos científicos y en PMApp [12].
3. *Log Processing*: En esta antepenúltima fase, tomando en cuenta la figura 2.3, el bloque de algoritmo seleccionado es el proceso de descubrimiento PALIA, el cuál crea modelos gráficos a partir de logs. Esta app, en particular, utiliza autómatas paralelos cronometrados (TPA) como principal formalismo, siendo probado con éxito en entornos médicos, siendo los autómatas modelos computacionales, que consisten en un conjunto de estados bien definidos, un estado inicial, un alfabeto de entrada y una función de transición. Además, como última cuestión a comentar, PMApp ofrece mucha flexibilidad al traducir TPA a otras notaciones como BPMN, siendo TPA una automatización del proceso de prueba, conectando los equipos con los instrumentos de prueba [12].

4. *TPA Processing*: En esta penúltima etapa, los procesadores de TPA calculan mapas específicos, estadísticas u otras modificaciones posteriores al descubrimiento del proceso. Bloques como el que figura en la figura 2.3, mantiene las posiciones de nodos y transiciones por si se presentan futuras ejecuciones [12].
5. *Rendering*: En esta etapa final, se aplica el proceso de resaltado de ideas, como, por ejemplo, mediante el bloque StatsInfoLabelTrantitionsBlock, que permite la customización de la información mostrada en las transiciones, que en este caso, es la duración media por rastro, pudiéndose elegir otras medidas antes de proceder a ejecutar el runner [12].

2.3 Aplicaciones usadas para la comparativa

En la anterior sección se comentaron aspectos relativos a lo que es el estándar XES y que compone al mismo, además de una breve justificación del porqué se seleccionó este estándar para llevar a cabo este trabajo, como también se destacaron las principales fases que posee la aplicación PMAApp. A continuación, se introduce la lista de aplicaciones basadas en minería de procesos que han sido seleccionadas para realizar una comparativa entre ellas. El objetivo de esta comparativa es identificar cuáles serían aquellos aspectos claves que debe de ofrecer una aplicación de minería de procesos para el análisis de datos clínicos.

Entre las aplicaciones que se usaron en la comparativa, destacan las siguientes:

En primer lugar, la denominada 'PMAApp', la cuál es una aplicación basada en minería de procesos desarrollada para ser utilizada específicamente en el ámbito clínico, y de la que cabe destacar los principales puntos fuertes, que son los siguientes:

1. *Flexibilidad*: Se pueden trabajar con múltiples fuentes de datos al mismo tiempo.
2. *Tamaño de carga*: No posee un límite de tamaño de archivo a importar, lo que es un punto muy fuerte, dejando sin posibilidad alguna la pérdida de parte de la información a analizar.
3. *Automatizable*: Permite configurar e interpretar procesos de datos.
4. *Customizable*: Se permite incorporar nuevos algoritmos, recursos y técnicas, lo que permite a los profesionales sanitarios mejorar los datos de análisis ofrecidos.
5. *Uso amigable para usuarios*: Interfaz gráfica bien diseñada y fácil de usar.
6. *Entendible*: Los modelos de los procesos ofrecidos son entendibles para los profesionales sanitarios.
7. *Transparente*: Se proporcionan métodos para evaluar la calidad de los datos durante la transformación y el filtrado de los mismos.

Debido a todo esto mencionado acerca de esta app, es el por qué se selecciona como mejor opción para los usuarios, selección que se detalla en el capítulo 3.

Ahora es el turno de la siguiente de las apps, la denominada 'Celonis', de lo que cabe comentar los siguientes aspectos:

Celonis es un software de minería de procesos que permite a los usuarios revisar y analizar las operaciones comerciales, con el fin de identificar posibles ineficiencias y administrar los flujos de trabajo. Este software ofrece análisis predictivo, el mapeo automatizado de procesos y el monitoreo en tiempo real, lo que ayuda a mejorar la productividad y optimización de sus procesos. Además, ofrece una amplia variedad de características, entre las que resaltan las siguientes [8]:

1. *Tamaño de carga*: Si que posee un límite de tamaño de archivo a importar, lo que es un punto desfavorable para esta app, ya que esto permite que parte de la información que se quiere analizar se pierda y no sea analizada.
2. *Flexibilidad*: Permite el trabajar con múltiples fuentes de datos al mismo tiempo.
3. *Uso fácil para el usuario*: La interfaz es sencilla de entender y de usar por parte de los usuarios.
4. *Transparente*: Se proporcionan opciones para poder medir la calidad de los datos ofrecidos durante los procesos de transformación y filtrado de los mismos.
5. *Entendible*: Todo lo que posee la app, incluyendo los resultados que ofrece la misma, es entendible por los consumidores finales, que en este caso, no solo pertenecen al ámbito sanitario.

En general, esta aplicación es bastante completa en el tema relativo a la minería de procesos, pero se enfoca mucho en el ofrecer características, por lo que está orientada a varios tipos de usuarios, no solo al ámbito comentado.

Ahora, con respecto a la siguiente aplicación denominada 'ProM', aunque no hay mucha información relativa a la misma, se puede resaltar lo siguiente [9]:

Es otra app basada en minería de procesos con un tiempo en el mercado bastante considerable, y con una interfaz gráfica que no es muy amigable para los usuarios finales, lo que hace también que no sea entendible. Por otra parte, la app no posee un límite de carga para los archivos a analizar, punto que resalta en este programa. Por último, tampoco posee métricas de tiempo ni abstracción gráfica, lo que hoy en día es fundamental que posean las apps basadas en minería de procesos, para poderle ofrecer una calidad de información mayor a los consumidores finales, ayudándolos así durante los momentos que requieran toma de decisiones.

Con esto comentado, se puede decir que es otra opción para los usuarios, pero que hoy en día existen otras opciones más completas y actualizadas, que ofrecen un mayor número de funcionalidades al usuario, y que permiten que los mismos tengan más libertad a la hora de manejar los distintos elementos que posee la interfaz gráfica.

Ahora es el turno de la siguiente app, denominada 'Disco', de la que cabe indicar lo siguiente:

Es otra de las apps basadas en minería de procesos, y, en este caso, no está específicamente diseñada para un ámbito en específico, sino que está diseñada para ser usada por diversos tipos de perfiles de personas según su área de trabajo. Entre los aspectos que destacan de esta app se encuentran [10]:

1. *Descubrimiento de procesos automatizado*: Es capaz de generar mapas de procesos automáticamente a partir de los datos crudos, siendo posible para los usuarios poder seleccionar el grado de abstracción y el tipo de visualización que desee el mismo.
2. *Interfaz fácil de usar*: Es simple de usar por parte de los usuarios, siendo para los mismos sencillo el tema relativo a encontrar los elementos que requieran en ciertos momentos para obtener un resultado en específico.
3. *Flexibilidad*: La app le da libertad a los consumidores finales para poder aplicar filtros a los resultados obtenidos, como también se puede trabajar con múltiples fuentes de datos al mismo tiempo.
4. *Animación de mapas de procesos*: La app ofrece una descripción general de los datos a partir de gráficos interactivos, con la posibilidad de profundizar en información detallada sobre cada actividad y recurso del proceso
5. *Manejo de proyectos*: Disco hace que sea más fácil y rápido crear copias de tus conjuntos de datos y filtrarlos para investigar aspectos específicos, lo que generará múltiples conjuntos de datos.

Por último, falta hablar acerca de la última de las apps comparadas, denominada 'Apromore', que al igual que la app anterior, no está diseñada para únicamente un ámbito, y de la que se pueden destacar los siguientes puntos [11]:

1. *Espacio de trabajo compartido de modelos y registros de procesos*: La app permite compartir fácilmente modelos de procesos y registros de eventos.
2. *Descubrimiento de mapas de procesos y modelos BPMN*: La app es capaz de generar un mapa de procesos del modelo BPMN directamente desde un log de eventos, con la posibilidad de cambiar entre vistas del proceso.
3. *Flexibilidad*: Los consumidores finales tienen la opción de aplicar distintos filtros a los resultados obtenidos, con el principal objetivo de acercarse más a lo que desean los mismos.

4. *Interfaz fácil de usar*: Es entendible para los consumidores finales, como también es sencillo el manejar los elementos que componen a la misma, reduciendo el tiempo en que estos consumidores consiguen el resultado esperado.
5. *Tamaño de carga*: Apromore posee un límite de carga para los archivos que se quieren analizar, lo que lleva a la posibilidad de perder información relevante y no obtener la salida con la calidad esperada.

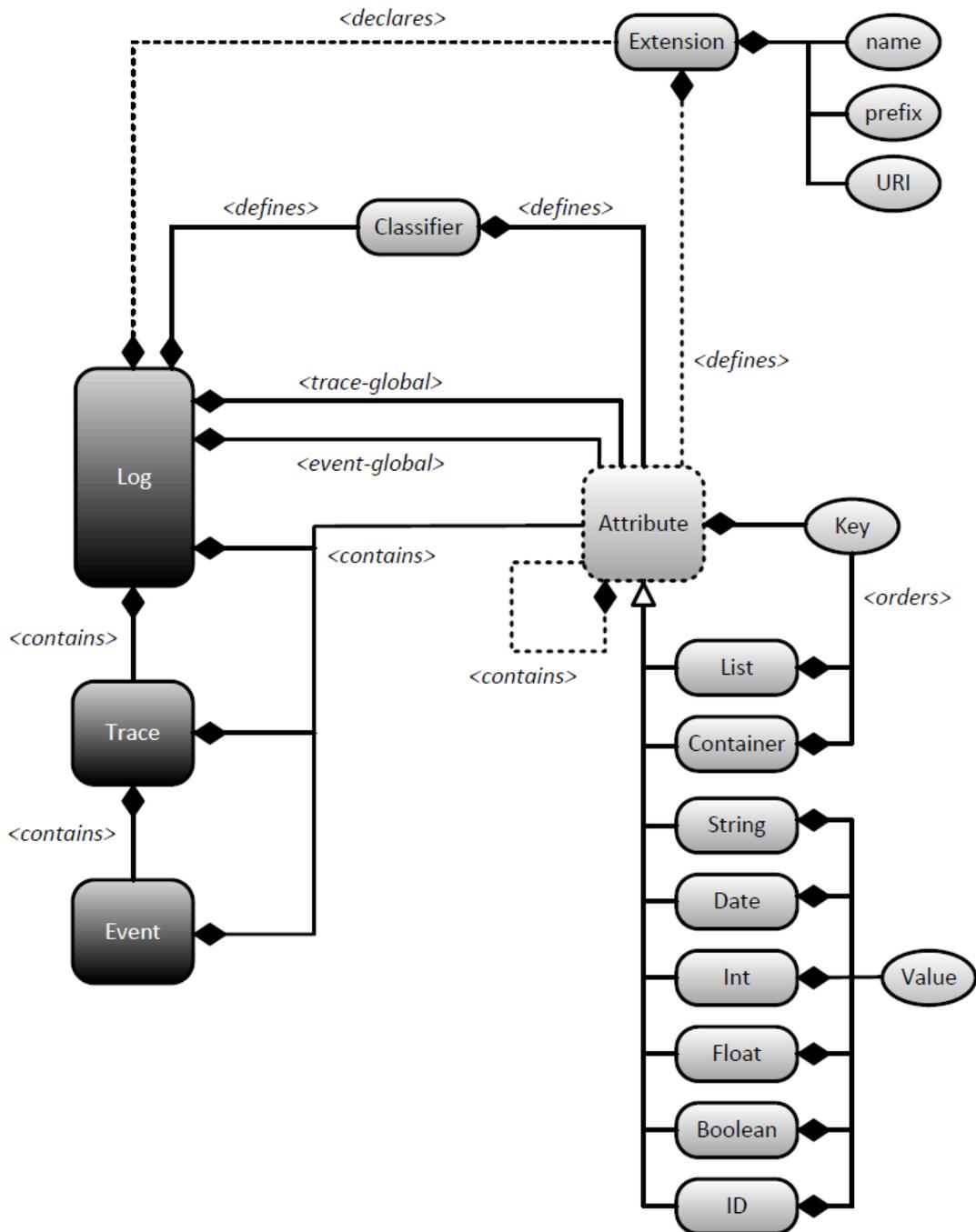


Figura 2.1: Diagrama de clases UML 2.0 para el metamodelo completo del estándar XES. Imagen proveniente del estándar XES [26].

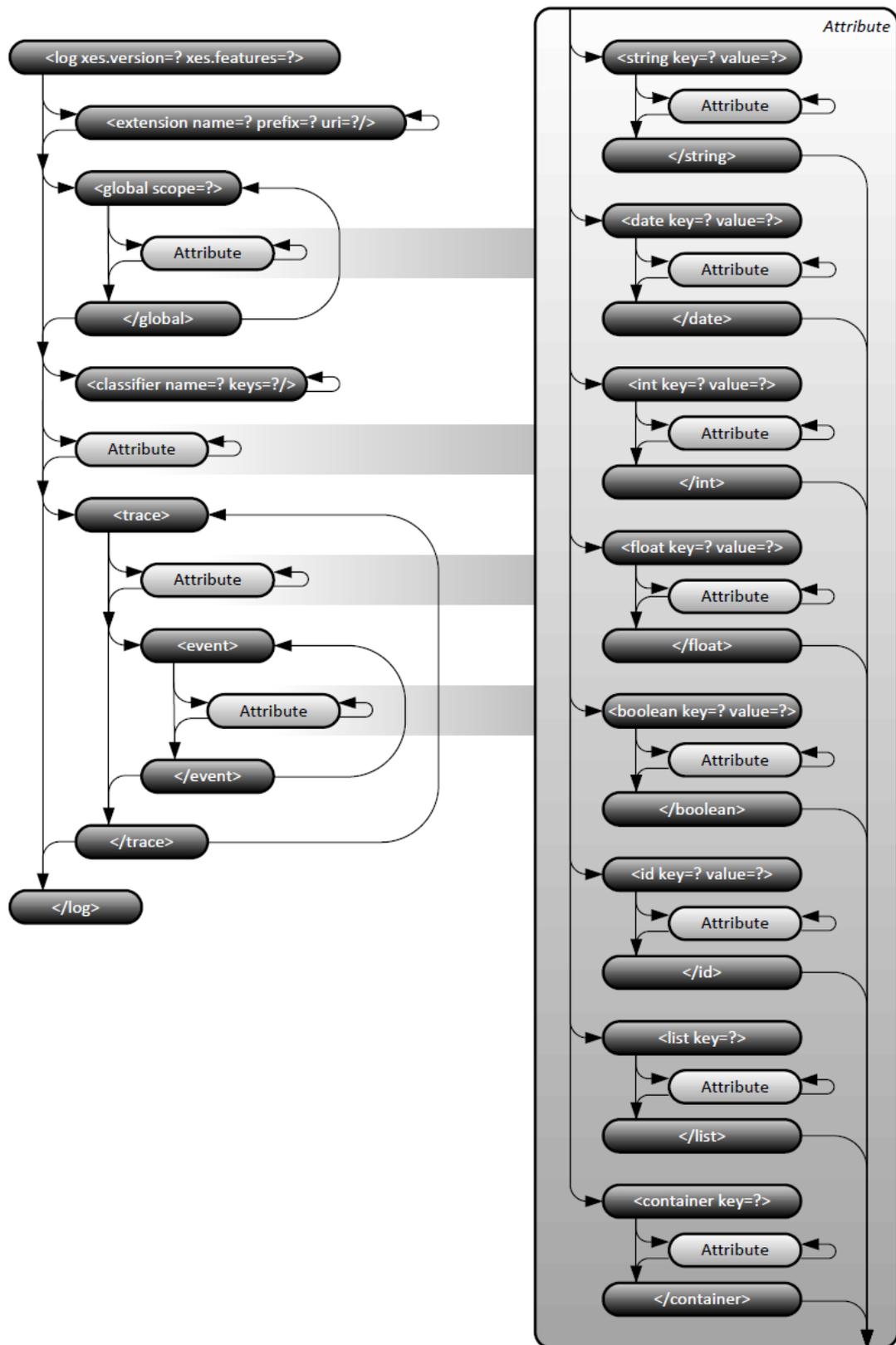


Figura 2.2: Diagrama de flujo de la máquina de estados para el estándar XES. Imagen proveniente del estándar XES [26].

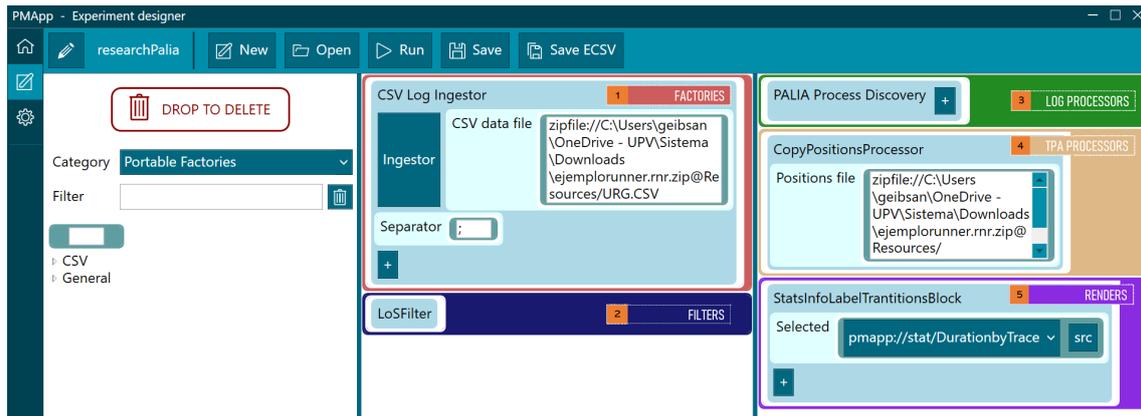


Figura 2.3: Fases previas al análisis de datos



Figura 2.4: Logo aplicación minería de procesos PMAApp



Figura 2.5: Logo aplicación minería de procesos Celonis



Figura 2.6: Logo aplicación minería de procesos ProM



Figura 2.7: Logo aplicación minería de procesos Disco

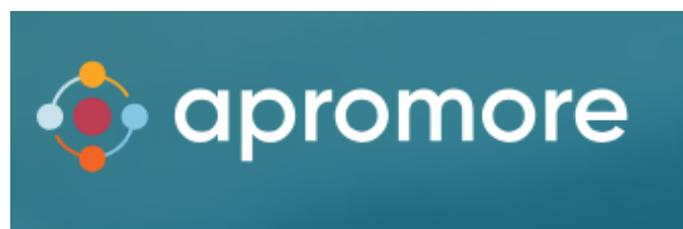


Figura 2.8: Logo aplicación minería de procesos Apromore

CAPÍTULO 3

Resultados

En el anterior capítulo se comentaron cuestiones relativas al estándar XES, entrando a detalle en temas relativos a qué es dicho estándar y qué componentes lo conforman, además de una breve argumentación de la elección de este estándar para el desarrollo de este trabajo, una breve descripción de qué es la minería de procesos y que fases la componen, y un breve resumen de las aplicaciones basadas en minería de procesos que existen hoy en el mercado, destacando las características que resaltan de las mismas.

3.1 Desarrollo módulos del estándar XES

Comenzando por los módulos desarrollados, es importante mencionar la estructura de clases presente y usada para llevar a cabo ese trabajo, que se puede visualizar en la figura 3.1, dónde cabe mencionar que esta estructura está diseñada para trabajar con archivos XES, teniendo como objetivo principal el incorporar funcionalidades de lectura, salvado y manejo de metadatos en formato XES.

Así pues, los módulos para la gestión del estándar XES corresponden con los objetivos *O1* y *O2* de la sección 1.2. Es importante destacar que este proyecto se realizó utilizando entornos como github, Visual Studio .NET, además de lenguajes de programación como C Sharp, cuestiones que se comentan a detalle en el capítulo 4.

Esto vino condicionado, por un lado, por la decisión de implementar dichos módulos para PMAApp, la cual se ha desarrollado en Visual Studio .NET C Sharp y donde los bloques desarrollados son de tipo factorías (capítulo 2).

En la figura 3.2, se puede ver la cabecera en el código, la cual hace que la librería (plugin) generada al compilar se convierta en un bloque listo para ser usado en PMAApp.

Además de lo ya comentado del módulo Reader, también es importante destacar el cómo se vería ese módulo en PMAApp, como también que clases fueron utilizadas para el desarrollo del mismo, lo que se puede visualizar en las figuras 3.3 y 3.4.

Donde en la figura 3.3 se puede ver la clase XesReader, implementada en la capa de lógica, en este caso, heredando de la clase PMLog, siendo ésta última una

clase que facilita el desarrollo del módulo al incorporar métodos que se pueden usar en la clase XesReader. Además, cabe destacar también que esta clase XesReaderBlock ya comentada también se utilizó durante el desarrollo del módulo reader, y que la misma fue implementada en la capa de presentación.

Por otro lado, ahora hablando con respecto al módulo saver, al igual que con el módulo reader, en este caso están presentes dos clases, las cuáles son la clase XesSaver (figura 3.5), que corresponde a la capa lógica, donde se traduce el LOG desde el formato de datos de PMApp a formato XES y se crea el archivo con toda la información correspondiente al mismo, y la clase XesTPASaver (figura 3.6), que hereda de BasicTPASaver, y que permite implementar esta funcionalidad relativa a la traducción de formato de datos.

En la figura 3.7, podemos observar que la pestaña 'PROCESS' tiene asociado un menú en la parte superior, donde se encuentra la opción de guardar (representado por un icono de diskette). Haciendo click sobre éste, se abre el menú modal 'Guardar como', desde donde podemos elegir el tipo de formato, estando disponible el formato 'XES (*.xes)'.

Ahora, en lo relativo al desarrollo de este TFG, y antes de mencionar una última cuestión relativa al módulo saver, cabe mencionar que el mismo se ha desarrollado dentro del grupo de investigación ITACA-SABIEN de la Universitat Politècnica de València¹, y donde el alumno se unió al equipo de trabajo de PMApp, y que, para el desarrollo de los módulos comentados, se realizaron varias reuniones. El alumno analizó el estándar XES, y lo presentó al equipo de desarrollo, con el objetivo de planificar el cómo se iba a llevar a cabo este proyecto, que fases lo iban a componer y el cómo se iban a desarrollar. Una vez iniciado el trabajo, se mantuvieron reuniones adicionales de control, con el fin de ir revisando el avance alcanzado.

3.2 Comparativa entre aplicaciones basadas en minería de procesos

Tras el desarrollo de los módulos del estándar XES para PMApp, se realizó una comparativa entre las aplicaciones basadas en minería de procesos elegidas. Esta comparativa corresponde con el objetivo O3 descrito en la sección 1.2 y donde se utilizaron 5 aplicaciones basadas en minería de procesos, las cuáles son: PMApp, Celonis, ProM, Disco y Apromore. Durante el desarrollo de la misma, se utilizaron diversos criterios de comparación, de los cuáles se hablará en detalle en la siguiente subsección.

3.2.1. Criterios de selección

Como se comenta en la anterior sección, a continuación se hablará a detalle acerca de los criterios de selección utilizados para realizar la comparativa, empezando por las cuestiones generales, donde se trataron los siguientes puntos:

¹ITACA-SABIEN: <https://www.sabien.upv.es/>

1. *Plataformas que soporta*: En qué tipos de sistemas operativos la aplicación es soportada, siendo estas plataformas, por ejemplo, Windows, Linux, Mac o basado en web.
2. *Versión de licencia*: Qué tipo de patente posee la aplicación, siendo una de estas patentes, por ejemplo, la de versión gratuita para una básica funcionalidad.
3. *Usuario final*: A qué tipo de consumidor final va dirigida la aplicación.
4. *Despliegue*: Qué tipo de lanzamiento posee la aplicación, siendo, por ejemplo, el tipo de lanzamiento como una aplicación de escritorio, o desde la nube.
5. *Automatización*: Si las aplicaciones permiten generar un archivo de configuración o no, siendo este archivo de configuración importante para almacenar parámetros y diversos ajustes necesarios para que la aplicación funcione correctamente, como también es importante ya que permite que el software de la aplicación se comporte de la manera esperada sin necesidad de modificar el código fuente.
6. *Capacidad de Personalización*: Si la aplicación les da libertad a los usuarios para agregar, modificar o eliminar cuestiones.
7. *Interfaz amigable para el usuario*: Si la app posee una interfaz que es fácil de usar por los usuarios.
8. *Entendimiento de la aplicación*: Si es intuitiva para los usuarios finales la app.
9. *Transparencia para el usuario*: Si los consumidores finales de la app no están al tanto de lo que sucede internamente en la misma, hablando, por ejemplo, de procesos internos. Es decir, si lo que sucede internamente es desconocido para el usuario final.

Lo que se obtuvo de las distintas aplicaciones comparadas mencionadas anteriormente, con respecto las cuestiones generales, se muestra en la figura 3.8:

Donde la Y significa que, sí que lo posee la aplicación, y N significa que no lo posee la misma.

Con respecto a las cuestiones relativas al manejo de la calidad de los datos, se consideraron los siguientes puntos:

1. *Tipos de archivos que permite importar la aplicación*: Que tipos de archivos se pueden analizar con la aplicación, como por ejemplo, de tipo .XES, .CSV, .SQL, entre otras.
2. *Conexiones a bases de datos*: Si la aplicación posee o no vinculaciones a bases de datos, como, por ejemplo, Oracle, JDBC, entre otras.
3. *Flexibilidad*: Si la app permite que se trabaje con múltiples fuentes de datos al mismo tiempo o no.

4. *Curación de los datos*: Si la app le informa al usuario como resuelve un problema que se presente durante la ejecución de la misma, como problemas relativos a como realiza la gestión de los datos que se pasen a la app para posteriormente ser analizados.
5. *Uso de técnicas de filtrado*: Si la app permite aplicar diversas condiciones/-restricciones sobre la salida que se le ofrece al usuario, como, por ejemplo, dividir un gráfico obtenido como salida por rangos de edad y género.

Lo que se obtuvo de las distintas aplicaciones comparadas mencionadas anteriormente, con respecto las cuestiones relativas al manejo de la calidad de los datos, se muestra en la figura 3.9:

Donde, nuevamente, la Y significa que la app sí que lo posee y la N significa que la app no lo posee.

Con respecto a las cuestiones relativas a las técnicas, se manejaron los siguientes puntos:

1. *Fase de descubrimiento del proceso*: Si la app posee la primera etapa fundamental en la minería de procesos, que implica extraer un modelo de proceso de datos de eventos registrados en un sistema de información, sin tener conocimiento previo explícito del proceso que generó esos datos.
2. *Verificación de Conformidad*: Si la app cumple con los estándares o no, como, por ejemplo, con el estándar de tiempos.
3. *Agrupamiento por trazas*: Si la app posee la funcionalidad de crear grupos de trazas, donde cada grupo tendría cierta información proveniente del archivo que se está analizando.
4. *Resaltado*: Si la app posee cuestiones relativas al resaltado de información relevante, como por ejemplo, el uso de colores para destacar cierta cuestión en un gráfico.
5. *Notación del Modelo de Salida*: Que formas de representar visualmente el modelo de proceso descubierto a partir de los logs de eventos posee la app.

Lo que se obtuvo de las distintas aplicaciones comparadas mencionadas anteriormente, con respecto las cuestiones relativas a las técnicas, se muestra en la figura 3.10:

Donde, nuevamente, la Y significa que la app sí que lo posee y la N significa que la app no lo posee.

Con respecto a las cuestiones relativas al análisis, se trataron los siguientes puntos:

1. *Buscar y filtrar gráfico*: Si la app permite localizar y aplicar diversas transformaciones al modelo del proceso obtenido como resultado.

2. *Tipos de atributos*: Si la app, al analizar el proceso y ofrecer el modelo de salida, incluye cuestiones como metadatos de traza y metadatos de eventos, teniendo en cuenta que los metadatos son datos que proporcionan información sobre uno o más aspectos de los datos.
3. *Métricas de tiempo*: Si la app posee medidas de tiempo que el usuario pueda seleccionar y aplicar o que ya de por sí se ofrezcan junto con el resultado, como, por ejemplo, la media, la varianza, el mínimo, el máximo, el sumatorio, la duración por traza, entre otras.
4. *Indicadores de rendimiento*: Si la app ofrece la opción al usuario de usar distintos tipos de gráficos para el análisis del modelo del proceso ofrecido, según lo que necesite obtener el mismo.
5. *Abstracción gráfica*: Si en la app es posible que el consumidor final combine distintos tipos de gráficos, por ejemplo, y obtenga resultados mucho más complejos.
6. *Granularidad alta*: Si el resultado que ofrece la app a los consumidores finales posee un nivel de detalle visual elevado.
7. *Valores anómalos*: Si la aplicación permite analizar los valores que no se ajustan a ningún estándar, es decir, que se desvían del comportamiento general de la mayoría de los valores.

Lo que se obtuvo de las distintas aplicaciones comparadas mencionadas anteriormente, con respecto a las cuestiones relativas al análisis, figura en la figura 3.11:

Donde, nuevamente, la Y significa que la app sí que lo posee y la N significa que la app no lo posee.

Ahora, como último punto, es importante también mencionar que criterios de comparación se trataron con respecto a las cuestiones relativas a la seguridad, los cuáles son:

1. *Acceso según tipo de usuario*: Si la aplicación, según el tipo de usuario que quiera utilizarla, ofrece ciertas funcionalidades, o para todos los tipos de usuarios se ofrecen las mismas.
2. *Tipo de autenticación*: Que medio/s de autenticación posee la app para poder empezar a usar la misma.
3. *Guardado de historial de acciones*: Si la aplicación posee alguna funcionalidad que permita guardar las acciones que se vayan ejecutando por cada usuario o no.

Lo obtenido de las distintas aplicaciones comparadas mencionadas anteriormente, con respecto a este último punto, se puede visualizar en la figura 3.12:

Donde, nuevamente, la Y significa que la app sí que lo posee y la N significa que la app no lo posee.

3.2.2. Metodología usada para comparación

En la anterior sección se comentaron a detalle los criterios de selección utilizados para desarrollar la comparativa, pero no se comentaron detalles acerca de que técnica de comparación se utilizó para llevar a cabo la misma, cuestión que se comenta a continuación.

En este caso, la comparativa se llevó a cabo mediante la técnica de Benchmarking, técnica que se utiliza para medir el rendimiento, en este caso, mediante distintos criterios de comparación, de las distintas aplicaciones basadas en minería de procesos, lo que permite a los usuarios/organizaciones tomar decisiones en función de sus necesidades y contexto, teniendo características detalladas de las distintas opciones que ayudan a los mismos a escoger la mejor opción. Esta técnica, en general, permite realizar una comparación de manera estructurada y sistemática que ayuda al proceso de toma de decisiones.

3.2.3. Resultados de la comparativa

En la sección anterior, se detallaron aspectos relativos a la metodología utilizada para elaborar la comparativa, y a continuación se mencionarán diversos asuntos relacionados con los resultados que se obtuvieron con la realización de esta comparativa, destacando principalmente las principales diferencias entre las aplicaciones comparadas, incluyendo varias capturas de pantalla de ciertas partes de la interfaz gráfica de cada una de las apps, y explicando cada una de ellas brevemente, para poder reforzar alguna de las características en qué estos programas se diferencian.

Empezando por la aplicación denominada PMAApp, cabe destacar lo siguiente. Con respecto a este programa destinada a la minería de procesos, cabe destacar de sus características generales, que es una aplicación que se puede utilizar

únicamente en el sistema operativo de Windows, por el momento, y que es una aplicación de tipo standalone, lo que quiere decir que es una aplicación de escritorio. Además, posee varios tipos de licencias, según para que se vaya a usar, entre las que destacan la versión gratuita, que posee un funcionamiento básico, y la versión comercial mucho más completa, y que es usada, en la mayoría de los casos, por usuarios pertenecientes al ámbito sanitario, como los médicos, por ejemplo. Hablando de otros aspectos más técnicos, es importante destacar que PMApp permite generar un archivo de configuración para cada una de las fases pertenecientes a la minería de procesos, lo que permite asegurarse de una ejecución eficiente. Además, PMApp es muy customizable, permitiendo a los usuarios añadir nuevos elementos al análisis, para hacer el mismo más sencillo, y que se adapte más a lo que en verdad quiere el usuario, como también su interfaz es bastante user-friendly.

Como punto siguiente, relativo al manejo de la calidad de los datos, cabe mencionar que PMApp permite importar archivos de tipo xes, csv, sql y ecsv, y estos archivos a importar no necesitan tener un tamaño máximo. También es importante mencionar que están presentes conexiones a bases de datos y que es muy flexible, permitiendo trabajar con múltiples fuentes de datos al mismo tiempo. Por último, con respecto a esta parte, y no menos importante, es importante no pasar desapercibido que PMApp también permite agregar filtros a los procesos representados, dándole mayores opciones a los usuarios de que puedan conseguir exactamente lo que quieren conseguir como resultado.

Siguiendo con este esquema, hablando ahora de los aspectos relativos a las técnicas, es importante destacar que PMApp posee la fase de descubrimiento de los procesos, lo cual es fundamental en aplicaciones de minería de procesos, como también posee verificación de conformidad, permitiendo asegurarse de que los procesos se ejecutan como es debido. Además, incluye cuestiones adicionales importantes, como el resaltar las cuestiones importantes en cada proceso representado, y el agrupar las trazas en clusters, lo que permite manejar la complejidad de los registros de eventos y la precisión de los modelos.

Ahora, como penúltimo punto, hablando de la parte relativa al análisis, cabe mencionar que esta aplicación posee métricas de tiempo, tablas y gráficas, y una granularidad alta que permite ver en detalle los procesos representados, incluyendo los datos anómalos, que, en este caso, al tratar datos médicos, son muy relevantes.

Como último punto, ahora hablando de cuestiones de seguridad, en esta aplicación, según el tipo de usuario, se podrán utilizar ciertas funcionalidades y otras no, y que, para poder empezar a usarla, es necesario autenticarse mediante un correo y una contraseña.

Una vez comentado lo anterior con respecto a esta app, a continuación se proporcionan diversas capturas de pantalla donde se pueden corroborar ciertas características que posee la misma y que se utilizaron en la tabla comparativa:

En estas dos primeras capturas de pantalla (Figuras 3.13, 3.14), figura primero la primera pantalla que ve un usuario al iniciar la aplicación, y segundo la pantalla previa a la ejecución del análisis del archivo, donde se puede observar que hay 5 fases, diferenciadas por colores, los cuáles son: Rojo, Azul, Verde, Amarillo y Morado, donde cada fase representa, respectivamente, la fase 'factories', la fase 'filtering', la fase 'LOG processing', la fase 'TPA processing' y la fase 'rendering', fases que se explican en el capítulo 2 a detalle. Además, también se pueden observar distintos tipos de extensiones, como la CSV, XES y la ECSV, cuestión que también figura como uno de los criterios de selección que forma parte de las cuestiones generales en la tabla comparativa. Como penúltimo aspecto, también es importante mencionar que la interfaz se nota que es entendible y fácil de usar, ya que, por ejemplo, se sabe dónde el usuario puede seleccionar la fase, que en este caso es en 'Category', que se puede incluir en cada una de las mismas, ya que al seleccionar una de ellas te salen automáticamente las opciones viables para la misma, si quieres aplicar algún filtro se sabe dónde se debe escribir el mismo, que en este caso es en 'Filter', debajo de 'Category', donde se le puede dar para iniciar el análisis del archivo, que en este caso es en 'Run', donde se puede salvar el experimento, donde se puede abrir un archivo, donde el usuario puede iniciar un nuevo experimento, entre otras cuestiones. Por último, cabe destacar que también en la aplicación se implementa el uso de colores para resaltar aspectos relevantes, característica que también está presente en la tabla comparativa y que posee esta aplicación.

Continuando con el tema relativo a las capturas de pantalla, con respecto a la primera de las siguientes dos imágenes (Figura 3.15), se puede observar cómo está presente un lector de bloques de archivos de tipo xes, como también un algoritmo de descubrimiento del proceso en la fase 3.

Ahora, con respecto a la segunda de las capturas (Figura 3.16), se puede visualizar como en la primera de las fases sigue estando presente el lector, pero en este caso, ya con la ruta del archivo xes que se quiere leer. Además, en la fase 4 se utiliza otro algoritmo para automatizar el proceso de prueba, es decir, modificaciones posteriores a la fase de descubrimiento del proceso.

Ahora, a continuación se proporcionan las últimas capturas de pantalla de la aplicación PMAApp, las cuáles ilustran como se ve un proceso una vez finalizado su análisis y toda la información relativa a éste, como también aspectos adicionales que se pueden incluir al proceso ya descubierto y analizado.

En las figuras 3.17, 3.18, 3.19, 3.20 se puede notar claramente como el proceso es descubierto y representado gráficamente, mediante colores, que en este caso representan si la actividad, por ejemplo, es un cuello de botella o no. Además, también se puede notar toda la libertad que tiene el usuario para poder agregar filtros, crear gráficos, modificar todo el tema de resaltado por colores, entre otros aspectos. Por último, y no menos importante, también cabe destacar que cada traza posee sus eventos, y que cada evento tiene sus características/metadatos propios.

Ahora, tomando en cuenta la segunda aplicación comparada, denominada 'Celonis', se puede resaltar lo siguiente:

Con respecto a esta otra aplicación, también destinada a la minería de procesos, cabe destacar de sus cuestiones generales, que es una aplicación que se puede utilizar en los sistemas operativos de Windows, Mac y basado en web a través de la nube, por el momento, y que la misma se puede ejecutar desde la nube o desde el escritorio de Windows y Mac. Además, posee varios tipos de licencias, según para que se vaya a usar, entre las que destacan la versión gratuita, que posee un funcionamiento básico, y la versión comercial mucho más completa, y que es usada, en la mayoría de los casos, por analistas de negocios y gestores de procesos. Hablando de otros aspectos más técnicos, es importante destacar que Celonis permite generar un archivo de configuración para cada una de las fases pertenecientes a la minería de procesos, lo que permite asegurarse de una ejecución eficiente. Además, Celonis también es muy customizable, permitiendo a los usuarios añadir nuevos elementos al análisis para hacer el mismo más sencillo y que se adapte más a lo que en verdad quiere el usuario, como también su interfaz es bastante user-friendly.

Como punto siguiente, relativo al manejo de la calidad de los datos, cabe mencionar que Celonis permite importar archivos de tipo xes, csv, xlsx, Google sheets, y estos archivos a importar no pueden superar el tamaño máximo, que, en este caso, es 1GB. También es importante mencionar que están presentes conexiones a bases de datos y que es muy flexible, permitiendo trabajar con múltiples fuentes de datos al mismo tiempo. Por último, con respecto a esta parte, y no menos importante, es importante no pasar desapercibido que Celonis también permite agregar filtros a los procesos representados, dándole mayores opciones al usuario de que puedan conseguir exactamente lo que quieren conseguir como resultado.

Siguiendo con este esquema, hablando ahora de los aspectos relativos a las técnicas, es importante destacar que Celonis posee la fase de descubrimiento de los procesos, lo cual es fundamental en aplicaciones de minería de procesos, como también posee verificación de conformidad, permitiendo asegurarse de que los procesos se ejecutan como es debido. Además, incluye cuestiones adicionales importantes, como el de agrupar las trazas en clusters, que permite manejar la complejidad de los registros de eventos y la precisión de los modelos, pero, en este caso, no se incluye el tema relativo a resaltar cuestiones importantes en los procesos representados, mediante el uso de colores, por ejemplo, lo que me parece también fundamental en aplicaciones de minería de procesos.

Ahora, como penúltimo punto, hablando de la parte relativa al análisis, cabe mencionar que esta aplicación posee métricas de tiempo, tablas y gráficas, y una granularidad alta que permite ver en detalle los procesos representados, pero, en este caso, hablando de Celonis, es importante recalcar que no se incluyen datos anómalos, sino que todo se ajusta a un estándar, lo que me parece una clara diferencia y desventaja, respecto a PMApp.

Como último punto, ahora hablando de cuestiones de seguridad, en esta aplicación, según el tipo de usuario, se podrán utilizar ciertas funcionalidades y otras no, y que, para poder empezar a usarla, es necesario autenticarse mediante un correo y una contraseña, y además de esto, está presente un doble factor de autenticación, diferencia también clara de Celonis con respecto a PMApp y Disco.

Una vez comentado lo anterior con respecto a esta app, a continuación se repite el proceso de las capturas de pantalla como con la app anterior:

En esta primera imagen (Figura 3.21) respecto a esta app, se puede notar claramente que existe un límite de carga para los archivos a analizar, en este caso 1GB, lo que es un punto débil de la misma, ya que, si el archivo es más grande que 1GB no se podrá analizar, por lo que los usuarios buscarán otra opción para analizar sus archivos que no posea un límite de carga.

En las figuras 3.22, 3.23 está presente información relativa al proceso, como los nombres de las distintas actividades, los orígenes de los eventos, los tiempos en que ocurrieron cada uno de los eventos, entre otros aspectos que figuran en la imágenes mencionadas.

En esta captura (Figura 3.24), como se comenta, se pueden visualizar las distintas fases por las que pasa el archivo durante su proceso de análisis, donde primero se transforman los datos, luego se crea el modelo del proceso y, por último, se carga dicho modelo.

A continuación, las siguientes 4 capturas:

En la primera de estas últimas 4 imágenes aportadas (Figuras 3.25, 3.26, 3.27, 3.28), si el usuario seleccionara la primera pregunta, el resultado que obtendría sería el que figura en la última de estas ilustraciones, donde se puede visualizar el proceso gráficamente, y donde también cabe resaltar que, en este caso, no se usan colores para destacar cuestiones importantes, lo que es otro punto débil para esta app ya que a los consumidores finales, en la mayoría de los casos, les interesa que lo más importante destaque de alguna manera, para poder detectar, por ejemplo, actividades que podrían ser un cuello de botella y que están retrasando al resto del proceso. Además, si el consumidor final seleccionara esta primera pregunta,

también obtendría otras informaciones, además de la representación gráfica del proceso, como las que figuran en la 2 y la 3 capturas de estas 4 últimas.

A continuación se detalla lo respectivo a la tercera de las apps comparadas, denominada 'ProM':

Con respecto a esta otra aplicación, también destinada a la minería de procesos, cabe destacar de sus características generales, que es una aplicación que se puede utilizar en los sistemas operativos de Windows, Mac y Linux, por el momento, y que la misma se puede ejecutar únicamente desde el escritorio de Windows, Linux y Mac. Además, cabe mencionar que, en este caso, el tipo de licencia que posee ProM es la GNU General Public License, lo que me parece un punto fuerte, ya que todo el mundo la puede usar con solo descargársela, sin necesidad de solicitar nada, y que es usada, en la mayoría de los casos, por analistas de negocios y gestores de procesos. Hablando de otros aspectos más técnicos, es importante destacar que ProM no permite generar un archivo de configuración para cada una de las fases pertenecientes a la minería de procesos, lo que es un punto desfavorable para esta aplicación, con respecto a PMAApp y Celonis, aplicaciones donde sí que se puede generar. Además, ProM también es muy customizable, permitiendo a los usuarios añadir nuevos elementos al análisis para hacer el mismo más sencillo y que se adapte más a lo que en verdad quiere el usuario, y en este caso, su interfaz no es user-friendly, otro punto desfavorable con respecto a esta app a comparación de PMAApp y Celonis, donde sus interfaces son user-friendly.

Como punto siguiente, relativo al manejo de la calidad de los datos, cabe mencionar que ProM permite importar archivos de tipo xes, csv, MXML, y estos archivos a importar, en este caso, no necesitan tener un tamaño máximo, lo que es un punto favorable para esta app, a comparación de Celonis, que sí que posee un tamaño máximo. También es importante mencionar que, en este caso, no están presentes conexiones a bases de datos, lo que es otro punto desfavorable para esta app, teniendo en cuenta como está todo lo relativo al análisis de datos actualmente, a comparación de Celonis y PMAApp, que sí que las tienen, como tampoco es flexible, por lo que no permite trabajar con múltiples fuentes de datos al mismo tiempo, cuestión que sí que pueden llevar a cabo PMAApp y Celonis. Por último, con respecto a esta parte, y no menos importante, es importante no pasar desapercibido que ProM también permite agregar filtros a los procesos representados, dándole mayores opciones al usuario de que puedan conseguir exactamente lo que quieren conseguir como resultado.

Siguiendo con este esquema, hablando ahora de los aspectos relativos a las técnicas, es importante destacar que ProM posee la fase de descubrimiento de los procesos, lo cual es fundamental en aplicaciones de minería de procesos, como también posee verificación de conformidad, permitiendo asegurarse de que los procesos se ejecutan como es debido. Además, incluye cuestiones adicionales importantes, como el de agrupar las trazas en clusters, que permite manejar la complejidad de los registros de eventos y la precisión de los modelos, y el tema relativo a resaltar cuestiones importantes en los procesos representados, mediante el uso de colores, por ejemplo, lo que me parece también fundamental en aplicaciones de minería de procesos para una mejor visualización de todo el proceso.

Ahora, como penúltimo punto, hablando de la parte relativa al análisis, cabe mencionar que esta aplicación, no posee métricas de tiempo, tablas y gráficas, cosas que sí que poseen PMAApp y Celonis, pero sí que posee una granularidad alta, que permite ver en detalle los procesos representados, incluyendo los datos anómalos, que, en este caso, al tratar datos me parece fundamental al hablar de minería de procesos.

Como último punto, ahora hablando de cuestiones de seguridad, en esta aplicación no hay ninguna cuestión relativa a la seguridad de esta, siendo estas cuestiones las siguientes: funcionalidades según tipo de usuario y métodos de autenticación para poder acceder a la app, lo que hoy en día es fundamental y que, sí que tienen PMAApp, Celonis y Apromore, por lo que este punto también es desfavorable para esta aplicación.

Una vez comentado lo anterior con respecto a esta app, a continuación se repite el proceso de las capturas de pantalla como con la app anterior, pero en este caso, incorporaré las primeras y comentaré aspectos relativos a las mismas, y lo mismo se hará con las otras dos:

Teniendo en cuenta estas dos primeras ilustraciones (Figuras 3.29, 3.30), se puede decir que la interfaz no es amigable ni fácil de usar para el usuario, ya que no hay ningún sitio que le diga al mismo donde podría cargar el archivo, ni donde clicar para poder ir a la fase de selección de algoritmo para analizar el archivo, ni como funciona el tema de la selección de los distintos algoritmos, lo que es un punto débil para esta app, y más hoy en día donde las aplicaciones que se desarrollan cada día se esfuerzan más por ofrecerle a los usuarios interfaces intuitivas.

Ahora, tomando en cuenta estas últimas dos percepciones (Figuras 3.31, 3.32) que figuran a continuación, también es importante mencionar que el resultado de lo más importante varía según la opción de representación gráfica que seleccione el usuario, lo que quiere decir que hay algunas representaciones gráficas del proceso que no incluyen resaltado, y hay otras que sí.

A continuación, se comenta la cuarta y penúltima app utilizada en la comparativa, denominada 'Disco':

Con respecto a esta aplicación destinada a la minería de procesos, cabe destacar de sus características generales, que es una aplicación que se puede utilizar en los sistemas operativos de Windows y Mac, por el momento, y que es una aplicación de tipo standalone, lo que quiere decir que es una aplicación de escritorio. Además, posee varios tipos de licencias, según para que se vaya a usar, entre las que destacan la versión gratuita, que posee un funcionamiento básico, y en este caso, muy limitado con respecto al tamaño de los archivos, y la versión comercial mucho más completa, y que es usada, en la mayoría de los casos, por analistas de negocios y gestores de procesos. Hablando de otros aspectos más técnicos, es importante destacar que Disco permite generar un archivo de configuración para cada una de las fases pertenecientes a la minería de procesos, lo que permite asegurarse de una ejecución eficiente. Además, Disco es muy customizable, permitiendo a los usuarios añadir nuevos elementos al análisis para hacer el mismo más sencillo y que se adapte más a lo que en verdad quiere el usuario, como también su interfaz es bastante user-friendly.

Como punto siguiente, relativo al manejo de la calidad de los datos, cabe mencionar que Disco permite importar archivos de tipo xes, csv, xls, mxml, xlsx, fxl disco log files, DSC Disco Project files, y estos archivos a importar, al estar usando la versión de prueba gratis, tienen un tamaño máximo que, en este caso, es muy limitante, cuestión que no posee PMAApp y ProM, por lo que es un punto desfavorable para esta aplicación. También es importante mencionar que están presentes conexiones a bases de datos y que es muy flexible, permitiendo trabajar con múltiples fuentes de datos al mismo tiempo. Por último, con respecto a esta parte, y no menos importante, es importante no pasar desapercibido que Disco también permite agregar filtros a los procesos representados, dándole mayores opciones al usuario de que puedan conseguir exactamente lo que quieren conseguir como resultado.

Siguiendo con este esquema, hablando ahora de los aspectos relativos a las técnicas, es importante destacar que Disco posee la fase de descubrimiento de los procesos, lo cual es fundamental en aplicaciones de minería de procesos, como también posee verificación de conformidad, permitiendo asegurarse de que los procesos se ejecutan como es debido. Además, incluye cuestiones adicionales importantes, como el resaltar las cuestiones importantes en cada proceso representado, y el agrupar las trazas en clusters, que permite manejar la complejidad de los registros de eventos y la precisión de los modelos.

Ahora, como penúltimo punto, hablando de la parte relativa al análisis, cabe mencionar que esta aplicación posee métricas de tiempo, tablas y gráficas, y una granularidad alta que permite ver en detalle los procesos representados, incluyendo los datos anómalos, que, en este caso, al tratar datos médicos, son muy relevantes.

Como último punto, ahora hablando de cuestiones de seguridad, en esta aplicación, no está presente el tema relativo a que, según el tipo de usuario, se podrán utilizar ciertas funcionalidades de la aplicación y otras que no, y que, para poder empezar a usarla, es necesario autenticarse mediante un correo electrónico y una contraseña.

Ahora, como se ha desarrollado con las anteriores apps, se incorporarán ciertas percepciones de Disco y se comentarán brevemente:

Teniendo en consideración las últimas tres percepciones (Figuras 3.33, 3.34, 3.35) con respecto a Disco, cabe destacar que la app si utiliza colores para resaltar aspectos relevantes, lo que hoy en día es fundamental en apps basadas en minería de procesos, ya que esto proporciona una ayuda extra a los usuarios en lo relativo a identificar posibles ineficiencias del proceso obtenido como resultado. Además, también se puede decir que la interfaz gráfica es bastante amigable para los consumidores finales, no siendo complicado el analizar un archivo, y teniendo todo muy detallado en base a lo que se quiera hacer.

Por último, en lo relativo a la última app utilizada en la comparativa, denominada 'Apromore', se puede mencionar lo siguiente:

Con respecto a esta otra aplicación, también destinada a la minería de procesos, cabe destacar de sus características generales, que es una aplicación que se puede utilizar en los sistemas operativos de Windows, Mac y Linux, por el mo-

mento, y que la misma se puede ejecutar desde la nube o desde el escritorio de Windows, Mac y Linux. Además, posee varios tipos de licencias, según para que se vaya a usar, entre las que destacan la versión gratuita, que posee un funcionamiento básico, y la versión comercial mucho más completa, y que es usada, en la mayoría de los casos, por analistas de negocios y gestores de procesos. Hablando de otros aspectos más técnicos, es importante destacar que Apromore, en este caso, que no estoy seguro de esto aún, no permite generar un archivo de configuración para cada una de las fases pertenecientes a la minería de procesos, lo que no permite asegurarse de una ejecución eficiente, lo cual es un punto desfavorable para esta aplicación, con respecto a PMAApp, Celonis y Disco, las cuáles sí que permiten generarlo. Además, Apromore también es muy customizable, permitiendo a los usuarios añadir nuevos elementos al análisis para hacer el mismo más sencillo y que se adapte más a lo que en verdad quiere el usuario, como también su interfaz es bastante user-friendly.

Como punto siguiente, relativo al manejo de la calidad de los datos, cabe mencionar que Apromore permite importar archivos de tipo bpmn, xes, csv, xlsx, parquet, zip, gz, y estos archivos a importar no pueden superar el tamaño máximo, que, en este caso, no está especificado, ya que solamente informa al usuario que el archivo a subir supera el límite permitido. También es importante mencionar que están presentes conexiones a bases de datos y que es muy flexible, permitiendo trabajar con múltiples fuentes de datos al mismo tiempo. Por último, con respecto a esta parte, y no menos importante, es importante no pasar desapercibido que Apromore también permite agregar filtros a los procesos representados, dándole mayores opciones al usuario de que puedan conseguir exactamente lo que quieren conseguir como resultado.

Siguiendo con este esquema, hablando ahora de los aspectos relativos a las técnicas, es importante destacar que Apromore posee la fase de descubrimiento de los procesos, lo cual es fundamental en aplicaciones de minería de procesos, como también posee verificación de conformidad, permitiendo asegurarse de que los procesos se ejecutan como es debido. Además, incluye cuestiones adicionales importantes, como el de agrupar las trazas en clusters, que permite manejar la complejidad de los registros de eventos y la precisión de los modelos, y, además, incluye el tema relativo a resaltar cuestiones importantes en los procesos representados, mediante el uso de colores en líneas que conectan actividades del proceso representado, o colores según la relevancia de cada actividad en el proceso, o debido a cuellos de botella, por ejemplo.

Ahora, como penúltimo punto, hablando de la parte relativa al análisis, cabe mencionar que esta aplicación posee métricas de tiempo, tablas y gráficas, y una granularidad alta que permite ver en detalle los procesos representados. Además de esto comentado anteriormente, también se pueden observar cuestiones (datos) anómalos que no se ajustan a ningún estándar, lo que, al estar tratando con datos médicos, es fundamental, ya que, en caso de que sea un paciente, es necesario saber porque lo que posee es un dato anómalo a diferencia del resto. Esto también podría ayudar a detectar nuevas mutaciones en alguna enfermedad, por ejemplo.

Como último punto, ahora hablando de cuestiones de seguridad, en esta aplicación, según el tipo de usuario, se podrán utilizar ciertas funcionalidades y otras no, y que, para poder empezar a usarla, es necesario autenticarse mediante un co-

reero y una contraseña, y además de esto, está presente un doble factor de autenticación, diferencia también clara de Apromore con respecto a PMAApp y Disco.

A continuación, se proporcionan ciertas capturas de pantalla de esta app, junto con una breve explicación, como se ha realizado con las 4 anteriores:

En esta primera percepción (Figura 3.36) se puede decir que la interfaz gráfica es bastante amigable para su uso por parte de los usuarios, estando todo muy claro y fácil de conseguir. Además, también se puede visualizar como el consumidor final, en caso de que lo desee, puede agregar filtros, cuestión que también figura en la comparativa como una característica de la app.

En estas dos siguientes percepciones (Figuras 3.37, 3.38) con respecto a estadísticas del proceso analizado, se pueden observar distintas informaciones del mismo, como el número de casos, el número de variantes y la duración de los mismos, donde cabe destacar que se usan métricas de tiempo, otra característica de 'Apromore' que también figura en la tabla comparativa, entre otras informaciones relevantes, como el tiempo en el que inicio cada caso y cada variante de éstos y cuando finalizaron los mismos.

Con respecto a estas dos últimas ilustraciones proporcionadas (Figuras 3.39, 3.40), es importante destacar que la app ofrece al usuario bastante flexibilidad a la hora de querer modificar algo relacionado con la salida gráfica obtenida, pudiendo, como ya se comentó, agregar filtros, cambiar la perspectiva de lo que se quiere visualizar y cambiar la notación al modelo BPMN. También se destacan los aspectos importantes mediante colores, lo que, como ya se resaltó antes, es fundamental.

3.2.4. Análisis de resultados

Una vez obtenidos los resultados de la comparativa, cabe destacar cuáles de las aplicaciones comparadas son las mejores opciones y por qué, lo que procede a continuación:

De todas ellas, PMAApp me parece la mejor opción, ya que es la única app de todas las comparadas que no se enfoca tanto en mostrar estadísticas, sino que está enfocada principalmente en el análisis de procesos relativos al ámbito sanitario, y en ofrecer un alto nivel de detalle de los mismos, permitiendo identificar todo lo óptimo y las ineficiencias presentes, permitiendo así poder realizar un análisis posterior al resultado y proceder a una optimización del proceso, si es el caso. Que no se enfoque tanto en ofrecernos estadísticas no quiere decir que no ofrece estadísticas, ya que si que las ofrece pero no en abundancia como el resto de aplicaciones, lo que, primero, hace que la interfaz gráfica esté colapsada de elementos, y aunque la app tenga una interfaz fácil de usar para los usuarios, el ofrecer y mostrar tantas estadísticas hace que la interfaz tenga demasiados elementos

al mismo tiempo, haciendo que los usuarios, aunque comprendan y sepan usar la interfaz de manera sencilla, tarden más tiempo en conseguir lo que buscan. Además, PMAApp me parece la mejor selección por los siguientes aspectos:

1. *Interfaz gráfica*: La interfaz gráfica me parece la mejor de todas, siendo sencilla, pero a la vez completa.
2. *Tamaño archivos*: El no tener límite de carga para los archivos que se quieren analizar es un punto muy bueno, además de tener diversos algoritmos como opciones para analizar el archivo y no perder ningún tipo de información.
3. *Filtros y Flexibilidad*: Se pueden aplicar una infinidad de filtros por parte de los usuarios, teniendo los mismos mucha libertad a la hora de querer modificar algún aspecto relativo a la salida obtenida y pudiendo conseguir de una forma óptima exactamente lo que quieren obtener.
4. *Granularidad*: El nivel de detalle es excepcional, permitiendo identificar en la salida obtenida cuellos de botella, algunas actividades del proceso que parece que no se están ejecutando como deben o que no están posicionadas donde deberían estarlo, lo que se consideran ineficiencias del proceso y, gracias a este nivel de detalle, se pueden optimizar todas ellas, permitiendo obtener una salida mucho más completa y óptima.
5. *Valores anómalos*: Gracias al punto anterior, además de poder optimizar el proceso, estas ineficiencias hay veces que se consideran valores que no se ajustan a la mayoría de los mismos, por lo que, gracias a esta profundidad de detalle, se pueden identificar y tratar.
6. *Acceso según usuario*: Aunque hay otras apps presentes en la comparativa que poseen doble factor de autenticación, PMAApp con el nivel básico de autenticación (usuario y contraseña) le basta, ya que, aunque no tenga doble factor, si que tiene un método de seguridad para poder empezar a usar la aplicación.

Las demás aplicaciones no quiere decir que sean pésimas opciones, sino que, en relación a el tema de analizar procesos relativos al ámbito sanitario, PMAApp es la indicada, siendo ésta la más orientada a ser usada por personal sanitario, sin ofrecer demasiadas estadísticas y enfocándose principalmente en los procesos.

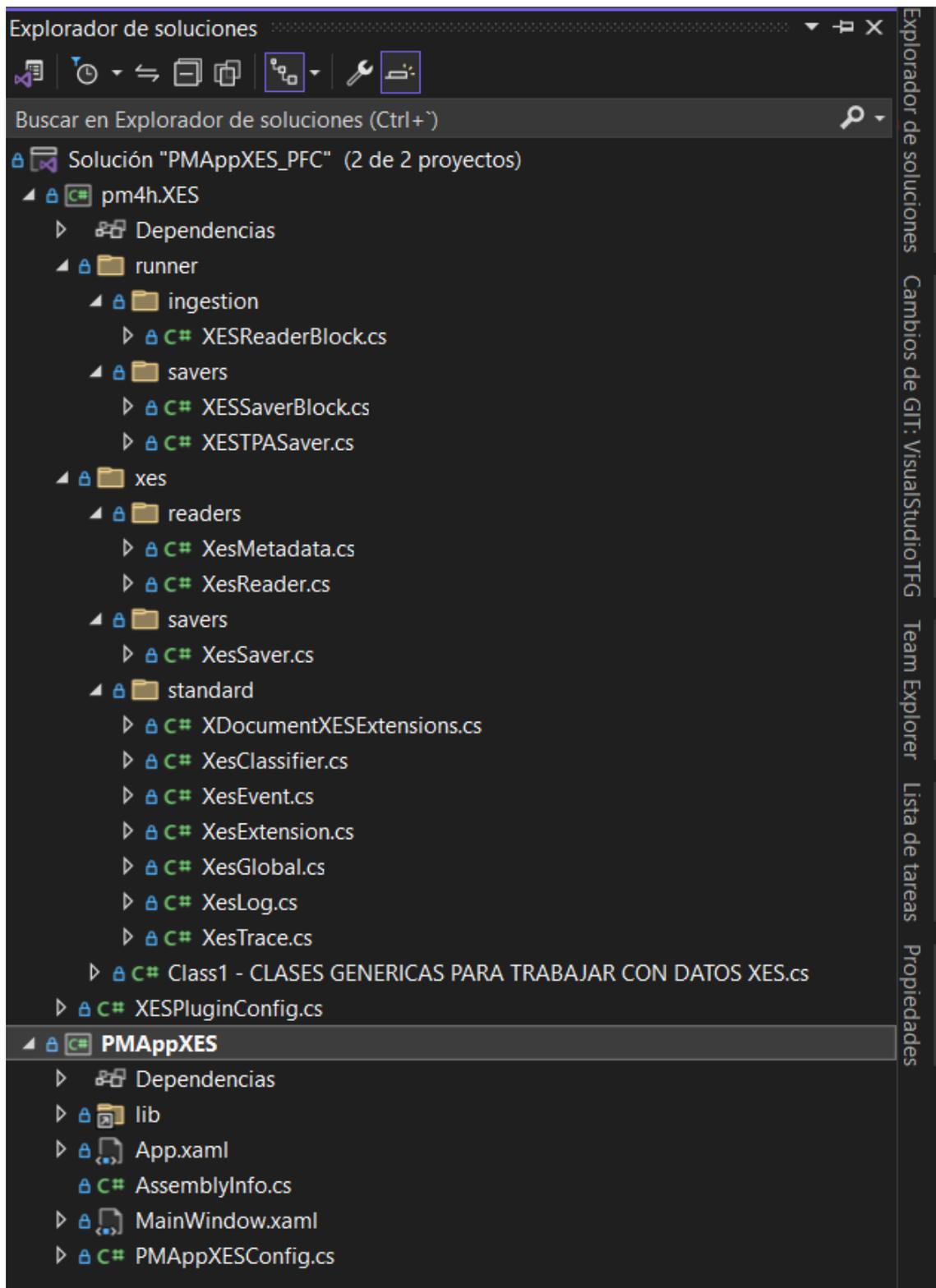


Figura 3.1: Estructura de clases

```

using pm4h.xes.standard;
using System;
using System.IO;
using System.Text;
using pm4h.xes.readers;
using pm4h.runner;
using pm4h.data;
using Sabien.Portable.Utils;
namespace pm4h.XES.runner.ingestion;

// La clase XESReaderBlock se encarga de leer los datos de entrada y pasarlos al XesReader de la carpeta standard para su procesamiento según todo lo que posee el estándar XES.
[RunnerElement(Name = "pm4h.XES.XesReader")]
0 referencias
public class XESReaderBlock : ILogCreatorActionRunner
{
    // Los datos de entrada que se van a procesar.
    //private string data;

    /// Constructor de la clase XESReaderBlock que recibe los datos de entrada.
    //public XESReaderBlock(string inputData)
    //{}

    // Asigna los datos de entrada pertenecientes al archivo de log que se pasa como argumento.
    // data = inputData;
    //}

    [RunnerExternalResourceFile(Default = "", Name = "XES data file",
        Description = "XES file containing the data to use at the experiment.",
        Domain = new string[] { "*.xes" }*)]
    1 referencia
    public IVirtualIO File { get; set; }
}

```

Figura 3.2: Parte del código desarrollado donde destaca la cabecera

```

using System;
using System.IO;
using System.Xml;
using Microsoft.CodeAnalysis.CSharp.Syntax;
using pm4h.data;
using pm4h.xes.standard;
using pm4h.XES.xes.readers;

namespace pm4h.xes.readers
{
    4 referencias
    public class XesReader : PMLog
    {
        1 referencia
        public void ReadFromXes(Stream data) // Lee los datos de entrada del documento XES y los procesa
        {

```

Figura 3.3: Clase usada para el desarrollo del módulo Reader

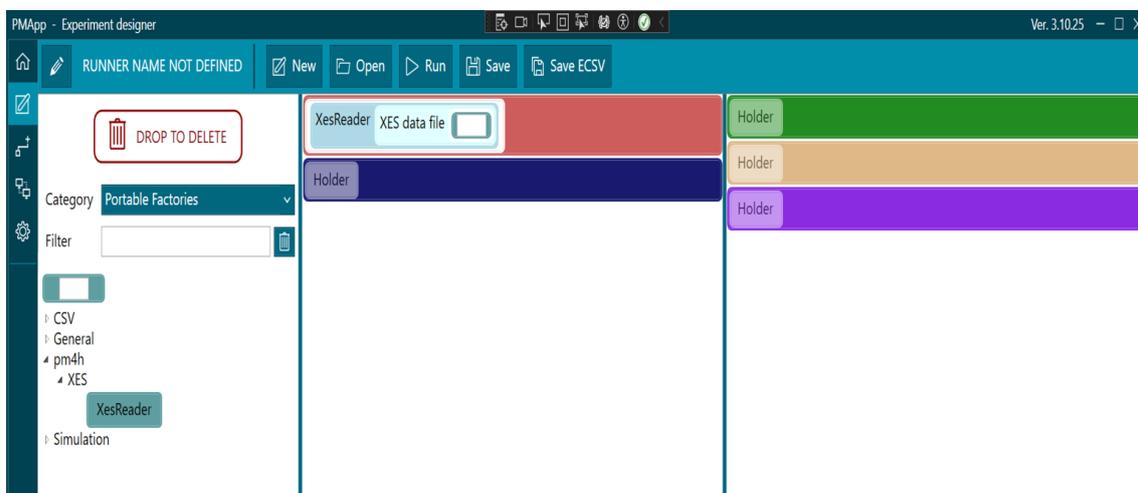


Figura 3.4: Visualización módulo Reader en PMAApp

```
using System;
using System.Collections.Generic;
using System.IO;
using System.Linq;
using System.Text;
using System.Xml;
using pm4h.data;
using pm4h.XES.xes.readers;

namespace pm4h.xes.savers
{
    2 referencias
    public class XesSaver
    {
```

Figura 3.5: Clase utilizada para desarrollar saver

```
using System;
using System.Collections.Generic;
using System.Diagnostics;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using pm4h.tpa.ipi;
using pm4h.utils.savers;
using System.Xml.Linq;
//using static pm4h.XES.xes.Class1__CLASES_GENERICAS_PARA_TRABAJAR_CON_DATOS_XES;
using pm4h.xes.savers;
using Accord.IO;

namespace pm4h.XES.runner.savers
{
    1 referencia
    public class XESTPASaver : BasicTPASaver //Esta clase tiene tres propiedades que son Name, Extension y Save que están definidas en la clase BasicTPASaver
    {
```

Figura 3.6: Otra clase utilizada para desarrollar el módulo saver

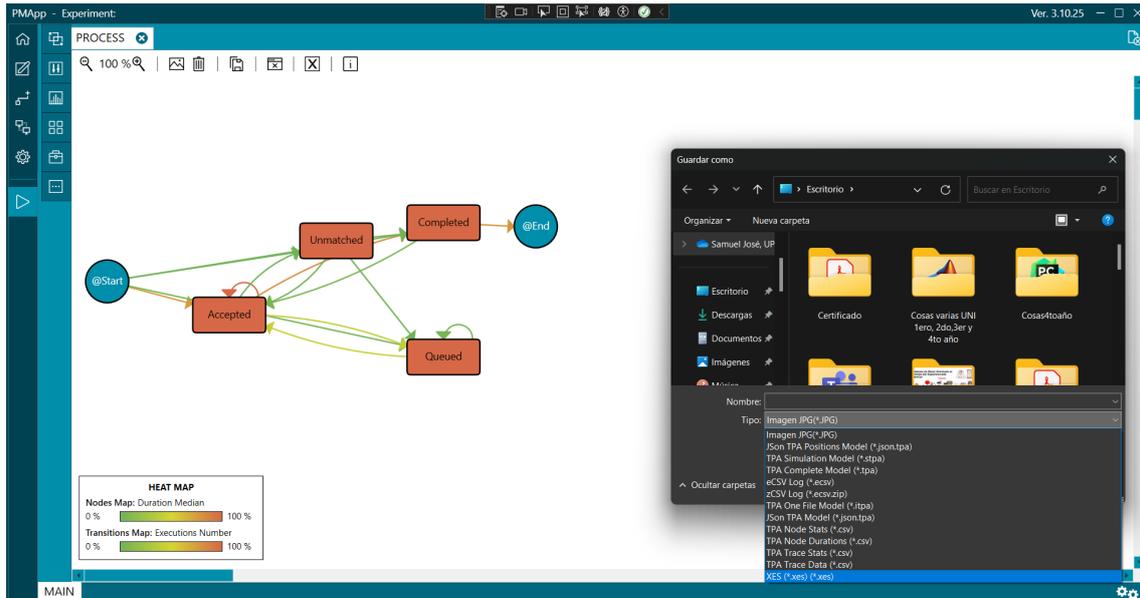


Figura 3.7: Visualización módulo Reader en PMAApp

FEATURE/TOOL	PMAApp	Celonis	ProM	Disco	Apromore
Supported platform	Windows for now	Nube/SaaS/basado en web,Windows, Mac	Windows, Mac, Linux	Windows, Mac	Windows, Mac, Linux
License	Other	Other	Other	Other	Other
License version	Free Version for a basic functionality, other licenses otorged by a commercial agreement, or a particular license for a single user	Free version for a basis functionality, companies licenses more complete and robust	GNU General Public License	Free version (demo), very limited for basic functionality, or commercial licences more complete and robust	Free version for a basis functionality, companies licenses more complete and robust
End user	Healthcare Professionals	Business analysts, Process managers and IT executives	Business analysts, Process managers and IT executives	IT executives and Healthcare Professionals	IT executives and Healthcare Professionals
Deployment	Standalone, as a desktop application	Hybrid --> Cloud and on-premise	Standalone	Standalone	Hybrid --> Cloud and on-premise
Automatizable	YES --> It is possible to generate a configuration file	YES --> It is possible to generate a configuration file	N	YES --> It is possible to generate a configuration file	N
Customizable	Y	Y	NO --> Because the interface is not very intuitive	Y	Y
User-friendly	Y	Y	intuitive	Y	Y
Understandable	Y	Y	N	Y	Y
Transparent	Y	Y	Y	Y	Y
GENERAL					

Figura 3.8: Resultados Cuestiones Generales Comparativa

FEATURE/TOOL	PMAApp	Celonis	ProM	Disco	Apromore
Import file types	.XES, .CSV, .SQL, .ECSV	.XES, .CSV, .XLSX, Google Sheets	.XES, .MXML, .CSV	.CSV, .XLS, .XLSX, .MXML, .XES, .FXL Disco log files, DSC Disco project files	.BPMN, .XES, .CSV, parquet, .XLSX, .ZIP, .GZ
Database connections	Y	YES --> SAP, ORACLE	N	Y	Y
Input size	Unlimited	Limited to 1GB	Unlimited	Unlimited	Limited (The specific limit does not appear)
Flexible	YES --> it is possible to work with more than only one data source at the same time	YES --> it is possible to work with more than only one data source at the same time	No --> it is not possible to work with more than only one data source at the same time	YES --> It is possible to work with more than only one data source at the same time, like data bases and log files	YES --> it is possible to work with more than only one data source at the same time, like data bases and log files
Data curation	Y	N	Y	N	N
Event log filtering	YES --> it is possible to use filter techniques	YES --> it is possible to use filter techniques	YES --> it is possible to use filter techniques	YES --> it is possible to use filter techniques	YES --> it is possible to use filter techniques
DATA QUALITY MANAGEMENT					

Figura 3.9: Resultados Cuestiones Manejo de calidad de Datos Comparativa

FEATURE/TOOL	PMAApp	Celonis	ProM	Disco	Apromore
Discovery	Y	Y	Y	Y	Y
Conformance checking	YES --> to ensure that processes are executed as planned	YES --> to ensure that processes are executed as planned	YES --> to ensure that processes are executed as planned	YES --> to ensure that processes are executed as planned	YES --> to ensure that processes are executed as planned
Trace clustering	YES --> having the possibility of making groups of traces	YES --> having the possibility of making groups of traces	YES --> having the possibility of making groups of traces	YES --> having the possibility of making groups of traces	YES --> having the possibility of making groups of traces
Enhancement	YES --> using things that stand out from the rest, like colors, for example	NO --> there are very few things that stand out from the rest	YES --> using things that stand out from the rest, like colors, for example	YES --> using things that stand out from the rest, like colors, for example	YES --> using things that stand out from the rest, like colors, for example
Output model notation	BPMN, WF	BPMN	BPMN	Petri Nets, Relation matrices, Variant Analysis and Conformance Checking	BPMN
TECHNIQUES					

Figura 3.10: Resultados Cuestiones Técnicas Comparativa

FEATURE/TOOL	PMAApp	Celonis	ProM	Disco	Apromore
Search and filter graph	Y	Y	Y	Y	Y
Attribute types	Event-level, Case-level	Event-level, Case-level	Event-level, Case-level	Event-level, Case-level	Event-level, Case-level
Time metrics	Average, Median, Variance, Min, Max, Deviation, Sumatory, Duration By Trace	Average, Median, Trimmed mean, throughout time analysis	NO --> it can be used by installing additional pluggins	Max repetitions, Total duration, Median duration, Mean duration, Max duration, Min duration	Average duration, Median duration, Max duration, Min duration
Charts and tables	YES --> having the possibility of using different types of graph types like a pie chart	YES --> having the possibility of using different types of graph types like a pie chart	NO --> it can be used by installing additional pluggins	YES --> having the possibility of using different types of graph types like a pie chart	YES --> having the possibility of using different types of graph types like a pie chart
Graph abstraction	Y	Y	NO --> it can be used by installing additional pluggins	N	Y
High granularity	Y	Y	YES --> for example, using the heuristics miner or the alpha miner	Y	Y
Outliers	YES --> having the possibility to see what is not normal or what do not adjust to the standard	NO --> All the details showed for each process are tight to the standard	YES --> having the possibility to see what is not normal or what do not adjust to the standard	YES --> having the possibility to see what is not normal or what do not adjust to the standard	YES --> having the possibility to see what is not normal or what do not adjust to the standard
ANALYSIS					

Figura 3.11: Resultados Cuestiones Análisis Comparativa

FEATURE/TOOL	PMAApp	Celonis	ProM	Disco	Apromore
Role-based access	YES --> depending of the user type, each user has specific functionalities	YES --> depending of the user type, each user has specific functionalities	NO --> ProM is focused on offering the tool	N	YES --> depending of the user type, each user has specific functionalities
User-authentication	Basic	Basic and 2FA (email)	NO --> ProM is focused on offering the tool	Basic	Basic and 2FA (email)
Audit logs	N	Y	N	N	N
SECURITY					

Figura 3.12: Resultados Cuestiones Seguridad Comparativa

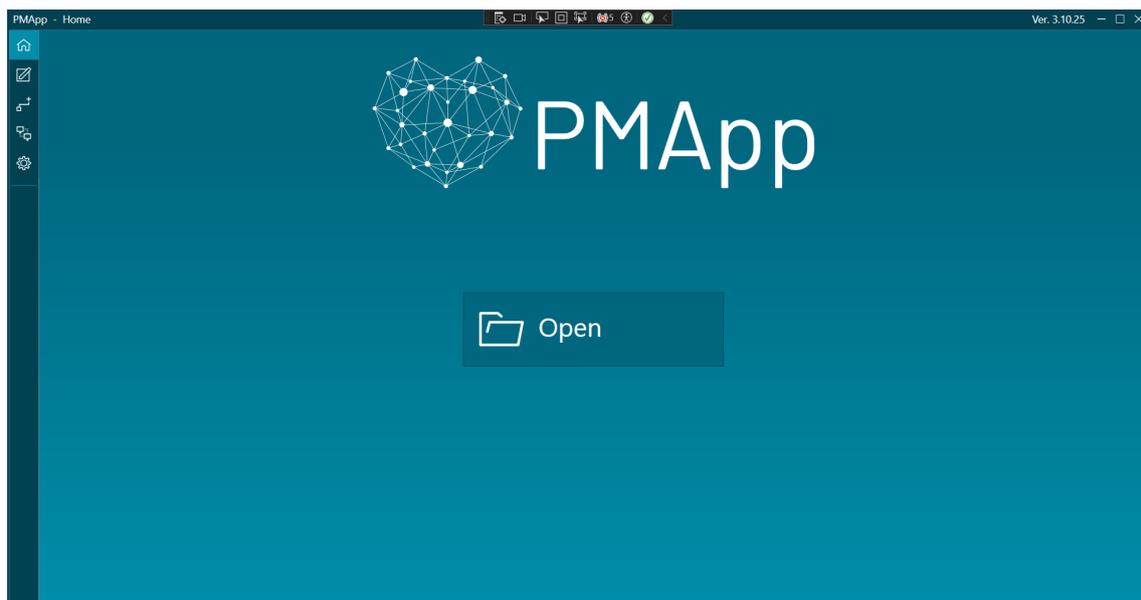


Figura 3.13: PMAApp - Pantalla inicial una vez se inicia la aplicación

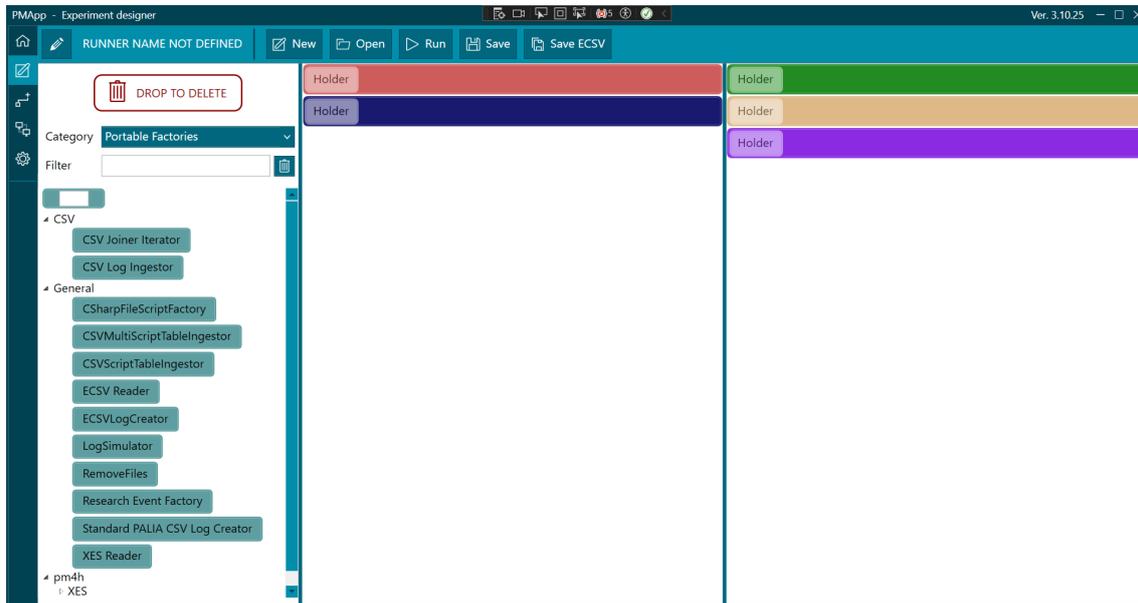


Figura 3.14: PMAApp - Pantalla previa a la ejecución del análisis del archivo

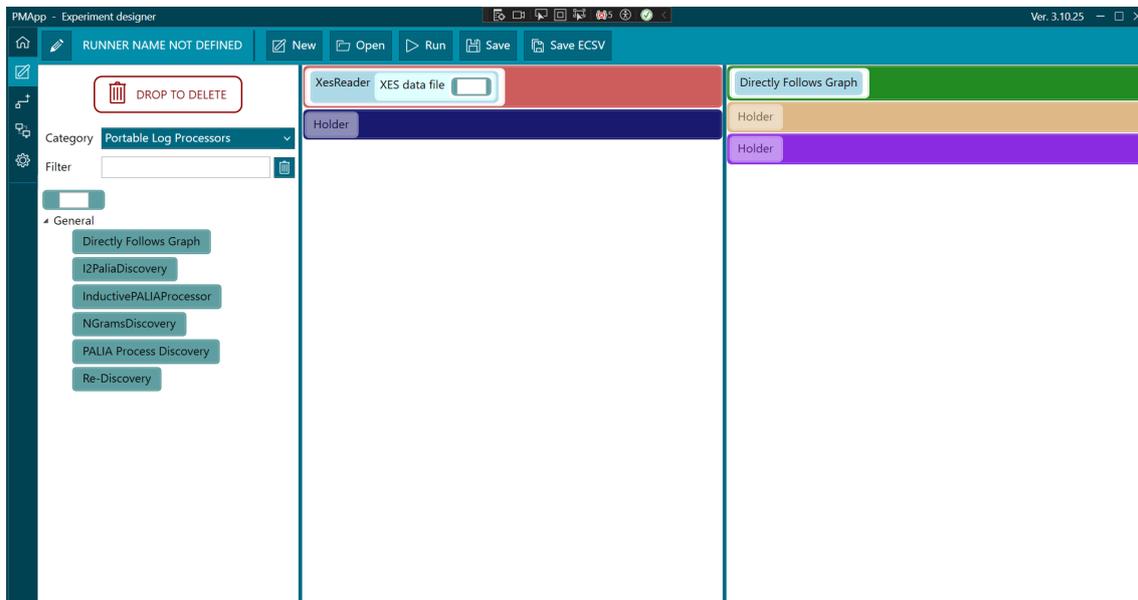


Figura 3.15: PMAApp - Pantalla con un lector de archivos xes en la fase 1 y un tipo de algoritmo de descubrimiento del proceso en la fase 3

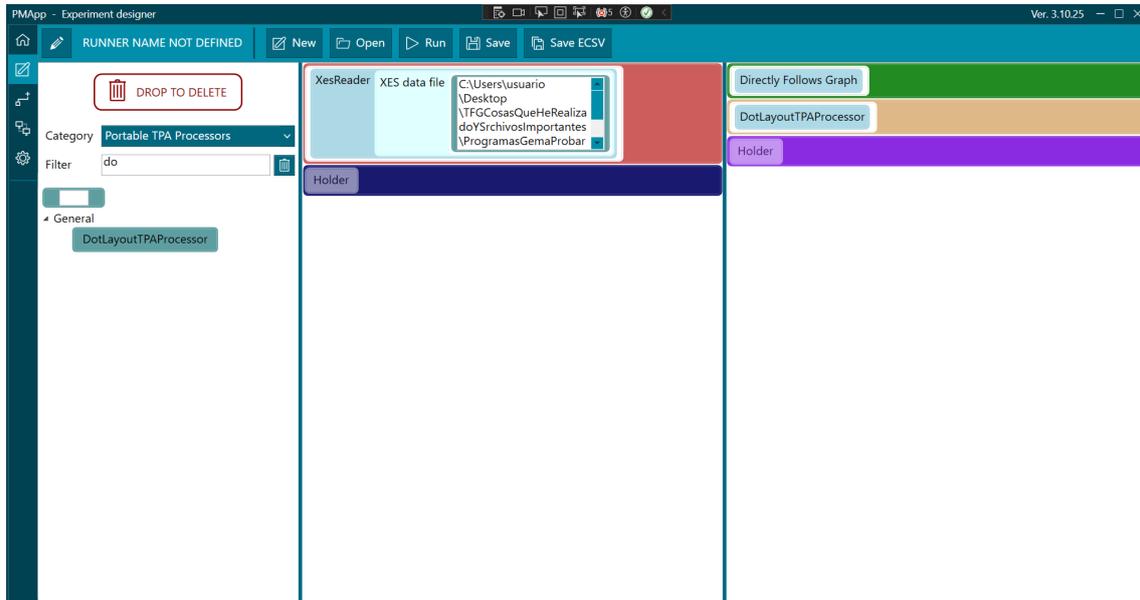


Figura 3.16: PMAApp - Pantalla con la ruta del archivo xes a leer, al algoritmo de descubrimiento del proceso, y otro algoritmo en la fase 4

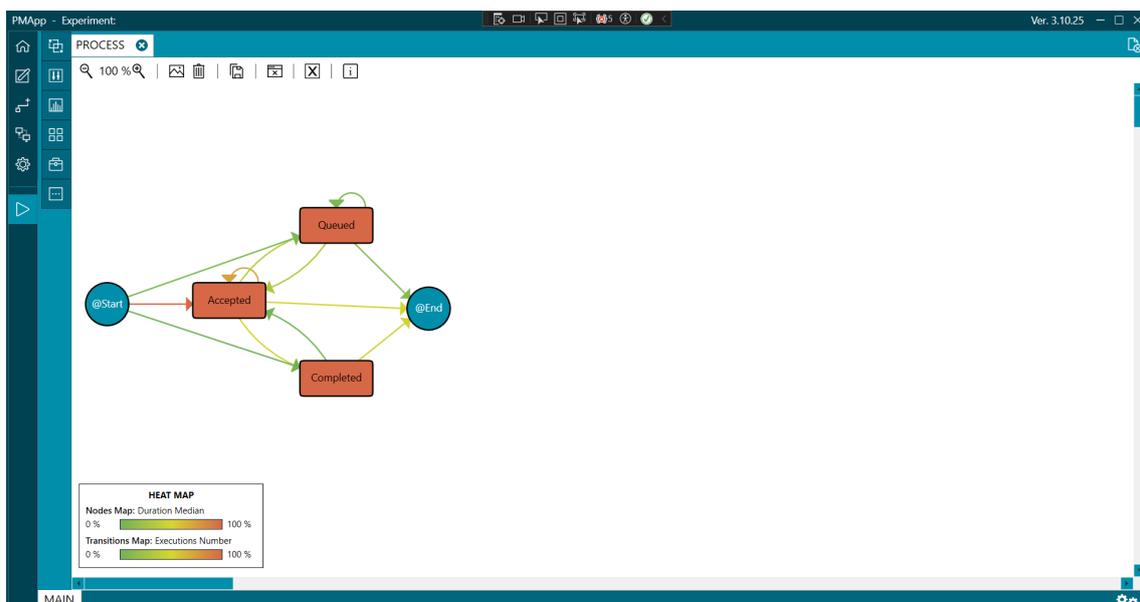


Figura 3.17: PMAApp - Pantalla donde figura el proceso descubierto

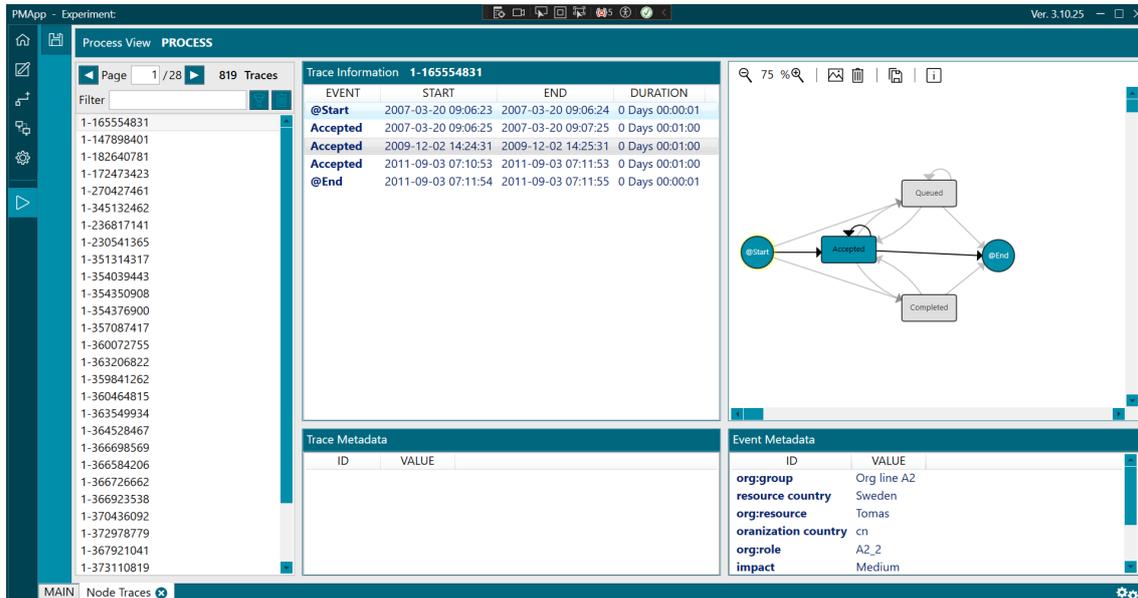


Figura 3.18: PMAApp - Pantalla donde figura información de las trazas del proceso descubierto

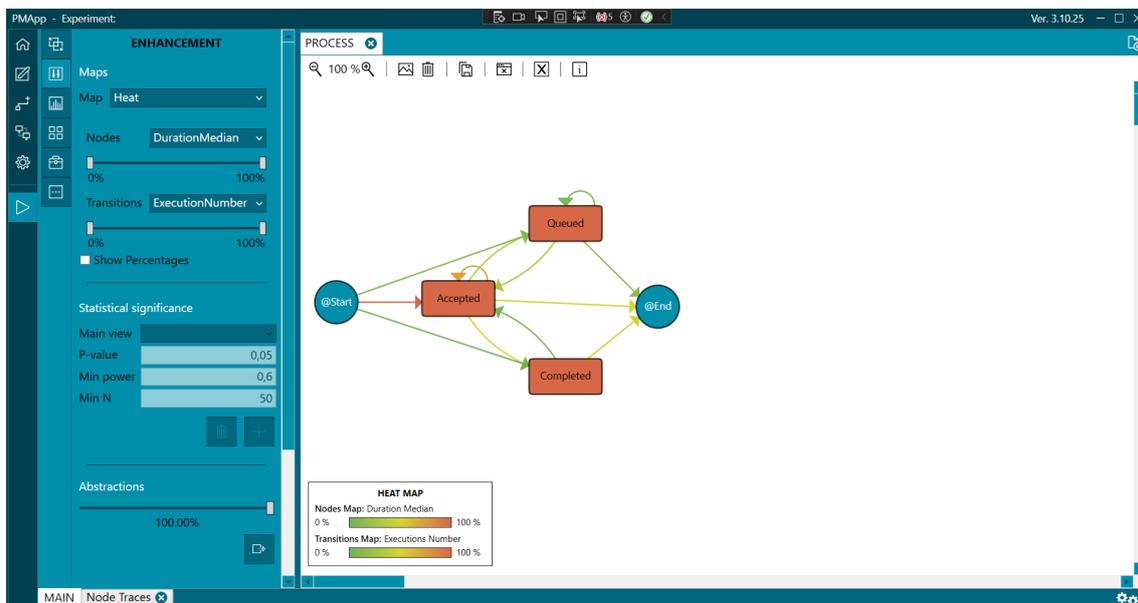


Figura 3.19: PMAApp - Pantalla donde figura información relativa a cuestiones que se pueden modificar con respecto al realce del proceso obtenido

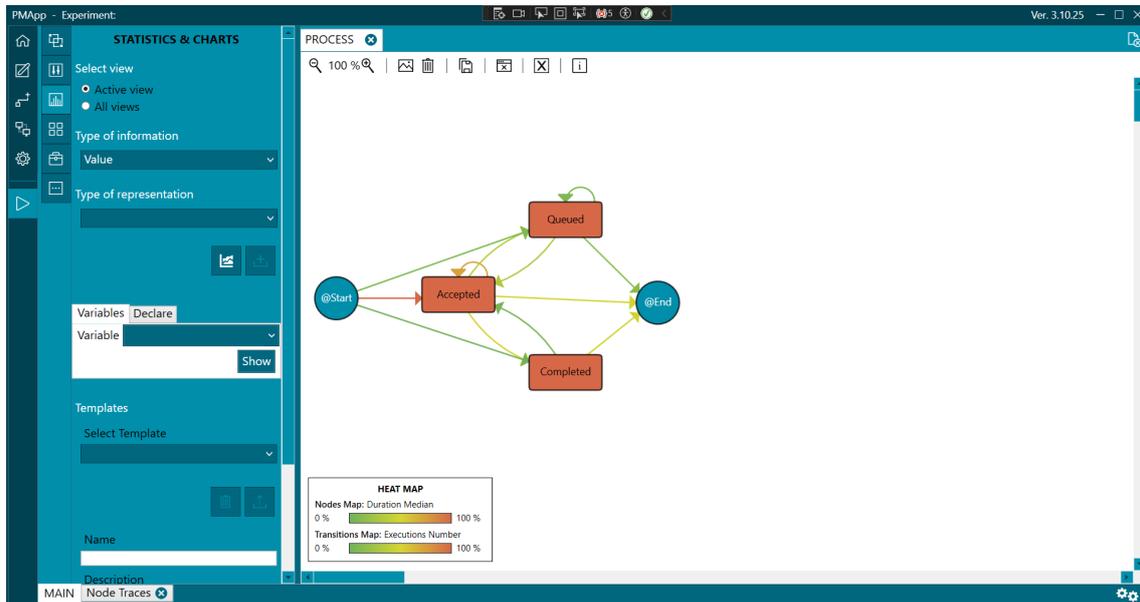


Figura 3.20: PMAApp - Pantalla donde figura información relativa a asuntos que se pueden modificar o agregar al proceso obtenido con respecto a estadísticas y gráficos

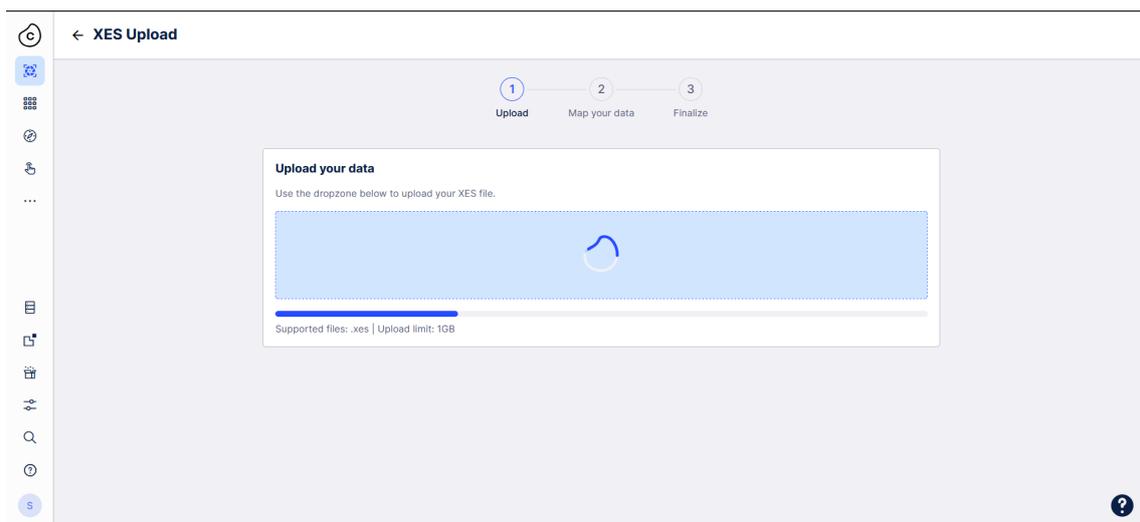


Figura 3.21: Celonis - Pantalla donde se carga el archivo xes para su análisis, siendo xes una extensión que acepta esta app

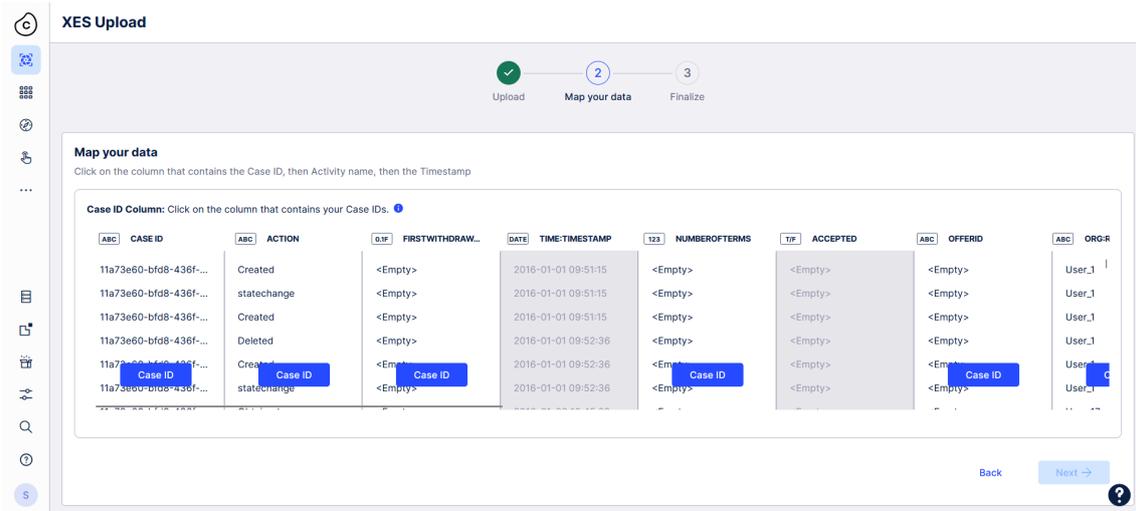


Figura 3.22: Celonis - Pantalla donde figura diversa información relativa al proceso

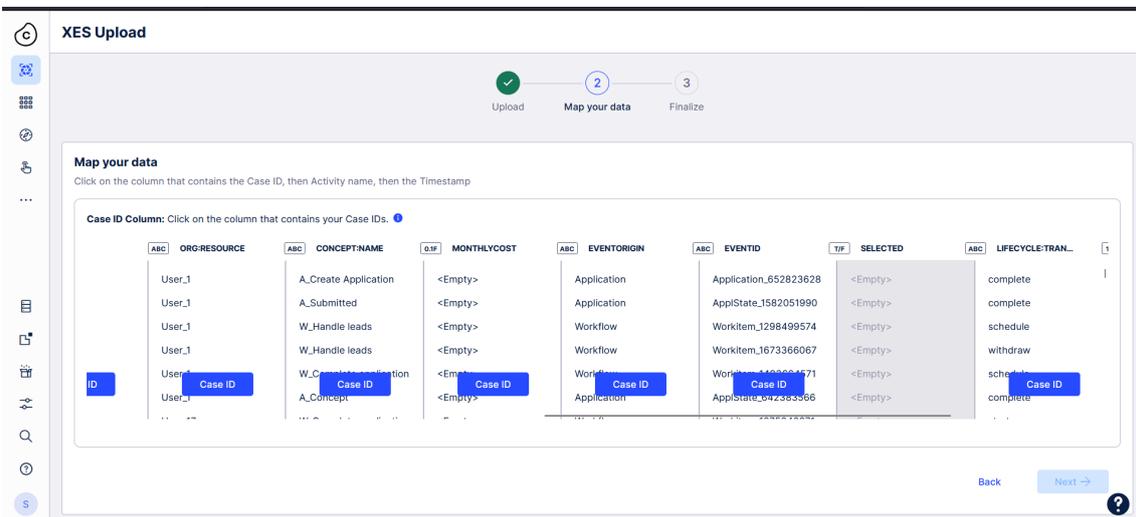


Figura 3.23: Celonis - Pantalla donde figura más información relativa al proceso

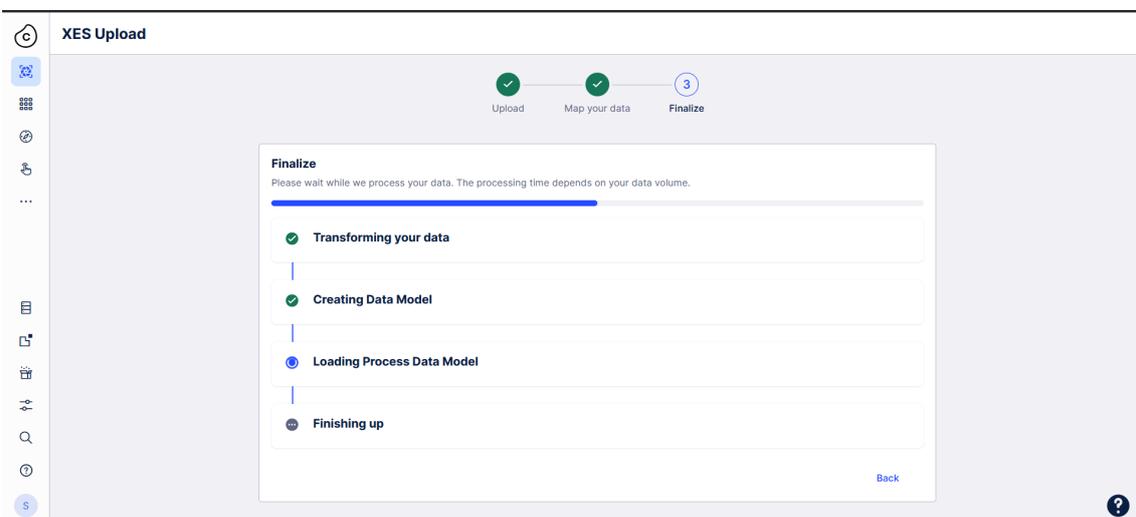


Figura 3.24: Celonis - Pantalla donde figuran las distintas etapas del análisis del archivo

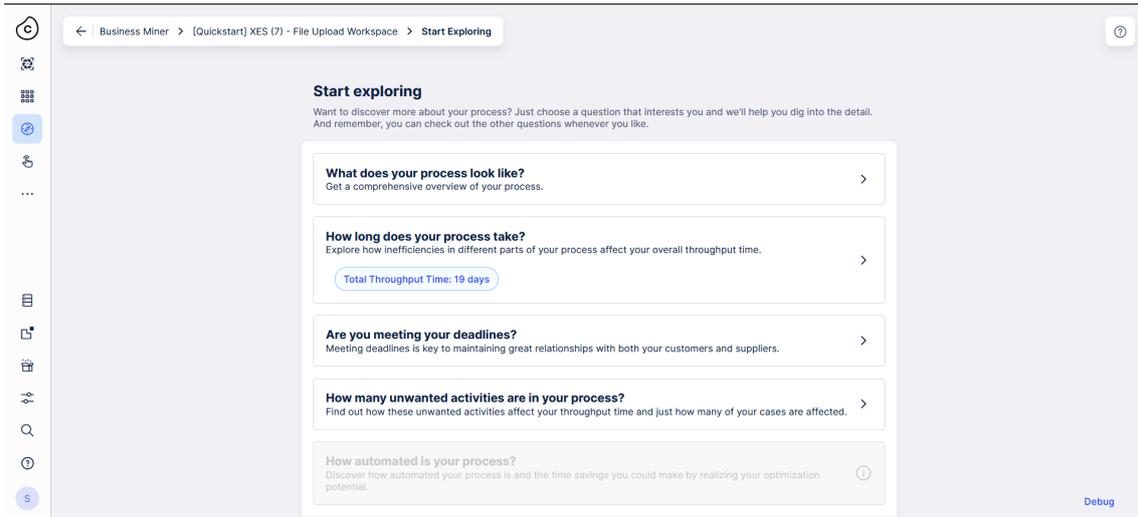


Figura 3.25: Celonis - Preguntas disponibles para el usuario

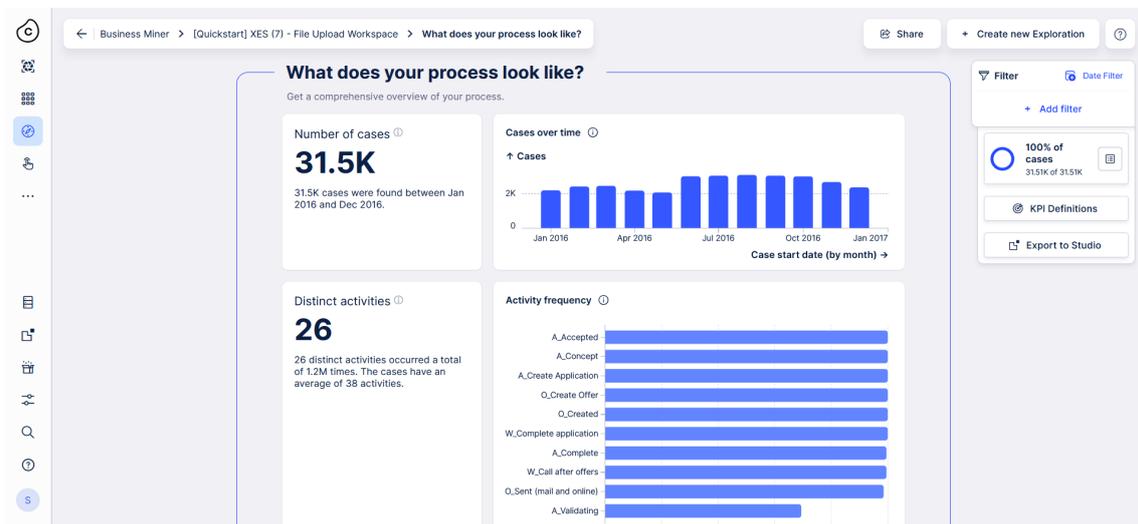


Figura 3.26: Celonis - Información relativa al proceso analizado

The screenshot shows a table of cases and a detailed view of activities for a specific case.

Case Id	# of Activities	Throughput
ffe5e61-...	32	22
ffe2c02-...	21	31
fff50707-...	22	31
fff36748-...	27	11
fff35439-...	28	7
fff30d4a-...	26	34
fff225f4-c...	21	5
ffef2fbf-b...	33	8
ffef2050-...	36	26

Activity	Time
A_Create Application	2016-10-31 12:42:59
W_Complete application	2016-10-31 12:42:59
W_Complete application	2016-10-31 12:42:59
A_Concept	2016-10-31 12:42:59
A_Accepted	2016-10-31 12:42:59

Figura 3.27: Celonis - Información adicional del proceso analizado

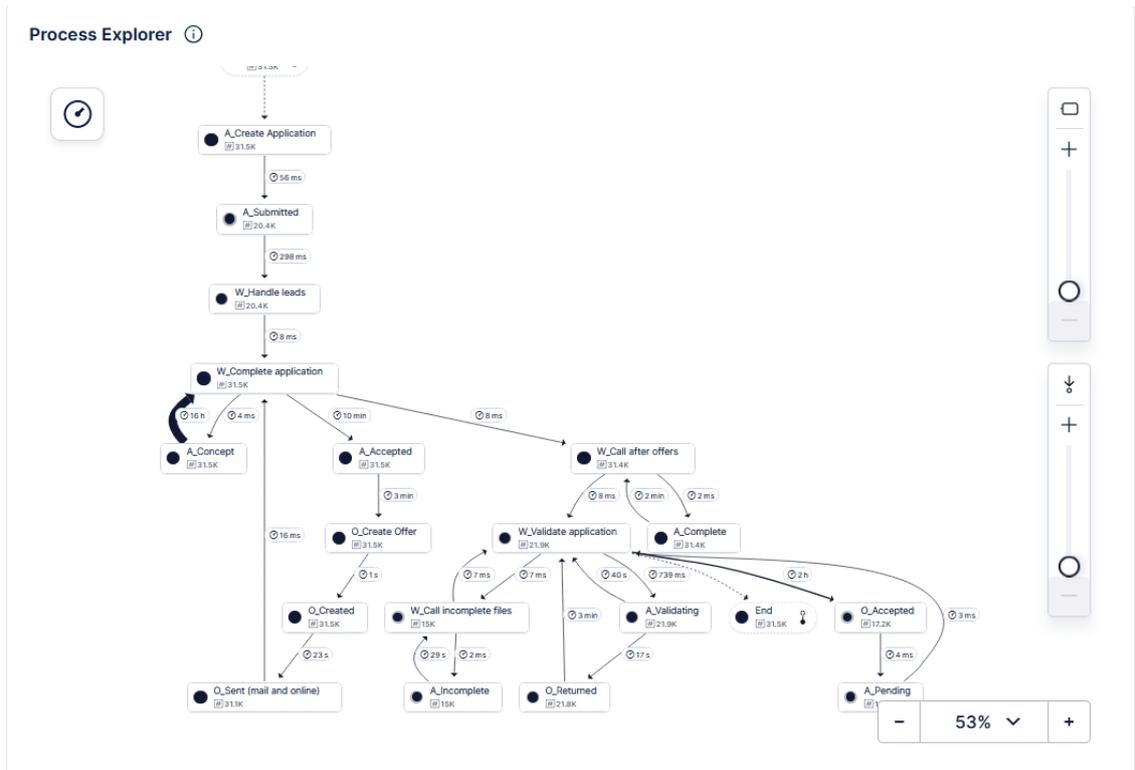


Figura 3.28: Celonis - representación gráfica del proceso

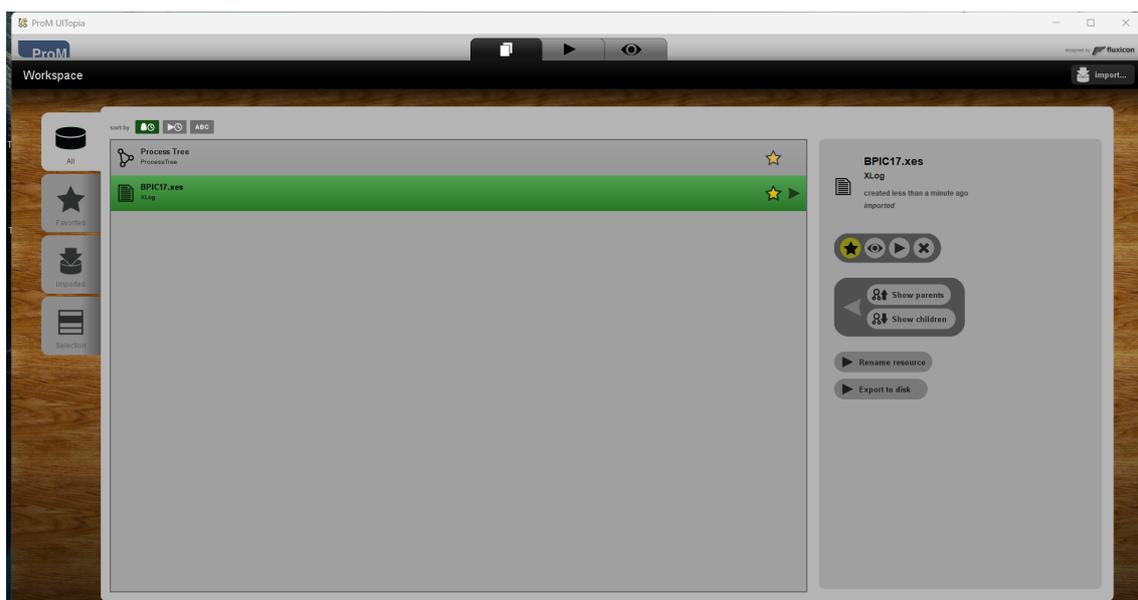


Figura 3.29: Prom - Pantalla inicial de la aplicación donde se cargaría el archivo para su posterior análisis

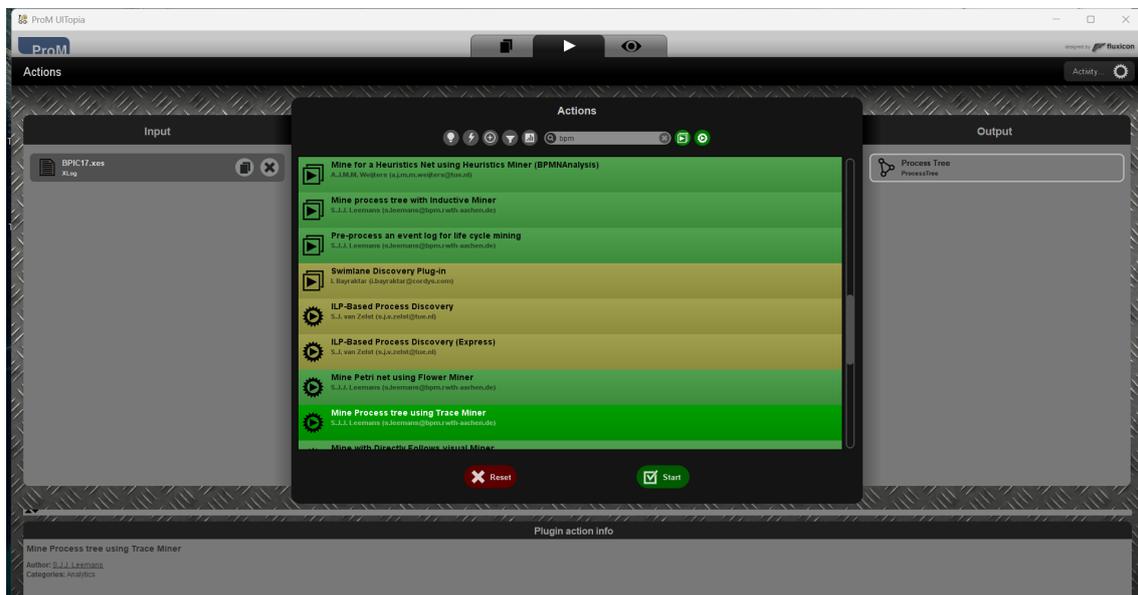


Figura 3.30: Prom - Distintas opciones de análisis del archivo cargado

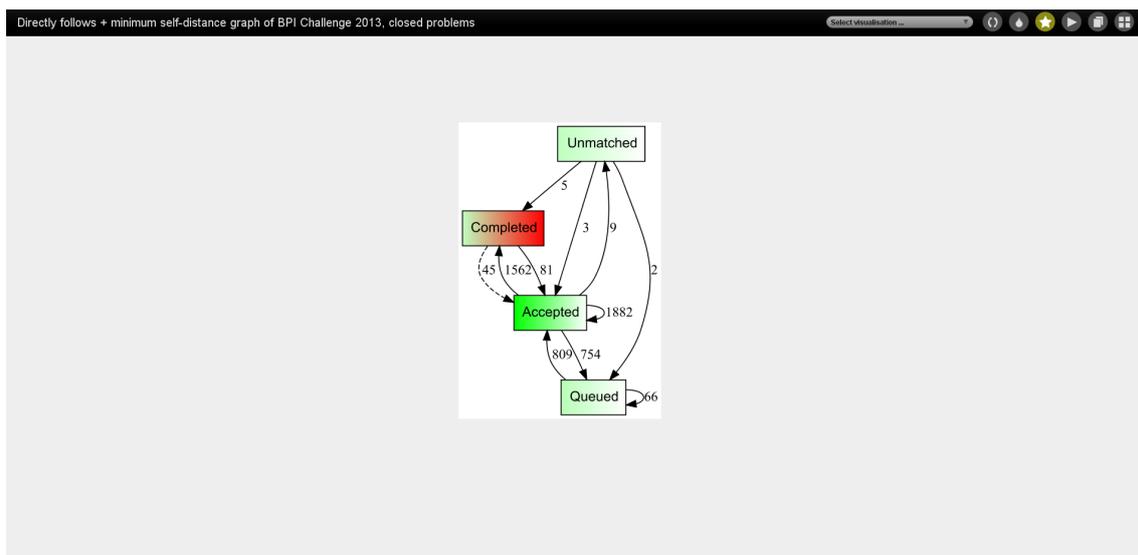


Figura 3.31: Prom - Salida gráficas del proceso analizado

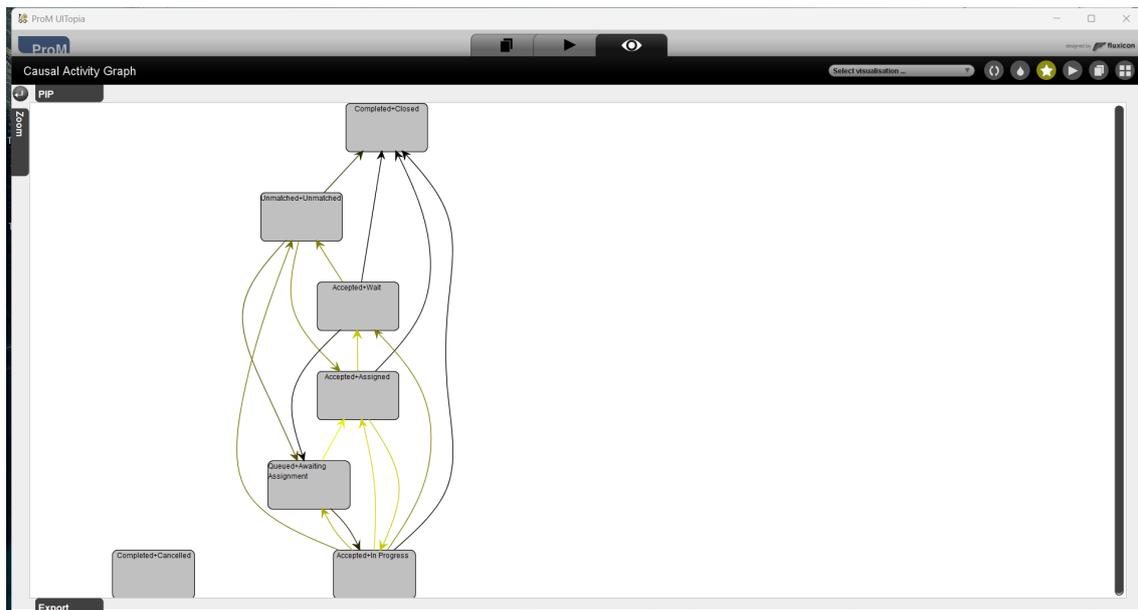


Figura 3.32: Prom - Otra posible salida gráfica del proceso analizado

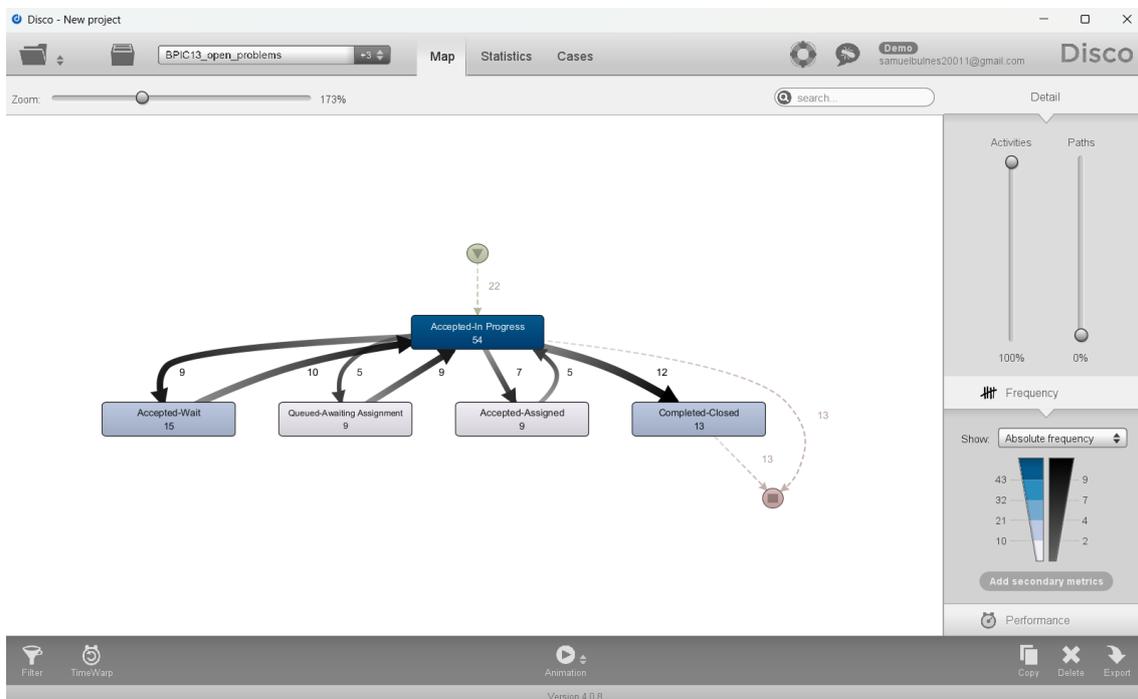


Figura 3.33: Disco - Imagen donde figura el proceso representado gráficamente

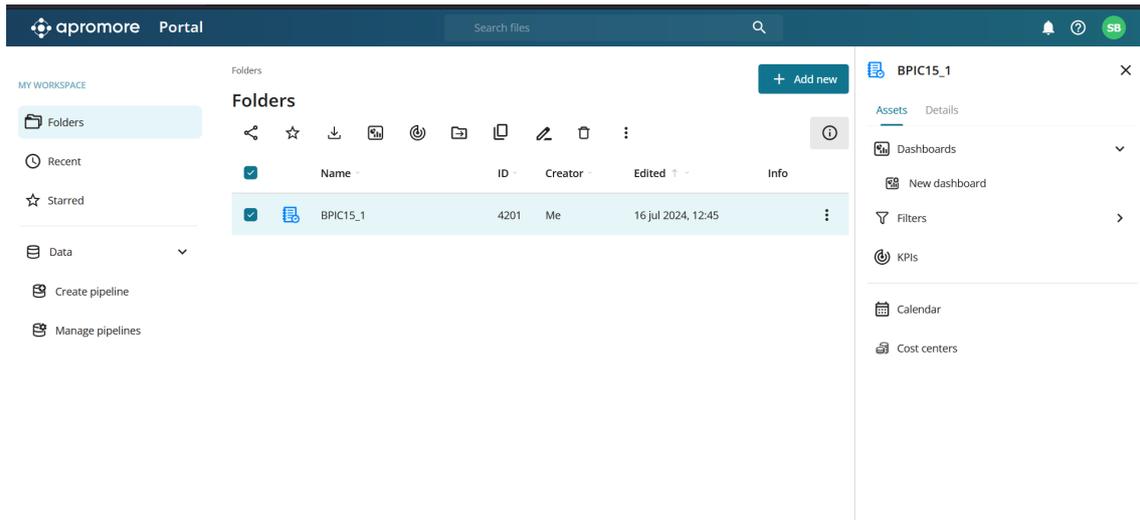


Figura 3.36: Apromore - Pantalla inicial donde se carga el archivo para su posterior análisis

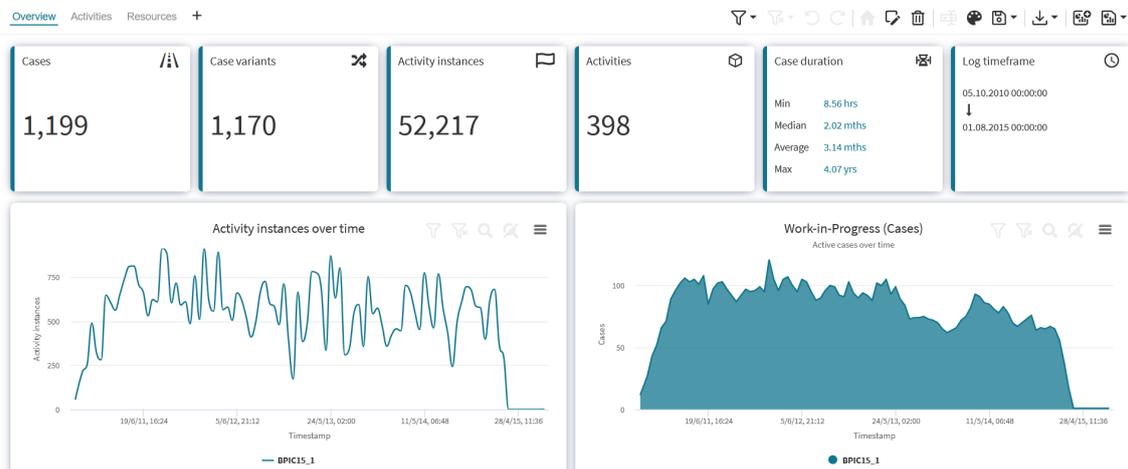


Figura 3.37: Apromore - Estadísticas del proceso obtenido como resultado

Case ID	Activity instances	Start time	End time	Duration
10009138	45	11/04/2014 00:00:00	05/06/2014 00:00:00	1.81 mths
10051383	57	17/04/2014 00:00:00	26/06/2014 00:00:00	2.3 mths
10053042	57	14/04/2014 00:00:00	03/07/2014 00:00:00	2.63 mths
10083315	58	17/04/2014 00:00:00	17/07/2014 00:00:00	2.99 mths
10093171	46	22/04/2014 00:00:00	10/07/2014 00:00:00	2.6 mths
10128431	56	25/04/2014 00:00:00	17/06/2014 15:34:15	1.76 mths
10153084	58	29/04/2014 00:00:00	17/06/2014 00:00:00	1.61 mths
10154600	47	30/04/2014 00:00:00	03/07/2014 00:00:00	2.1 mths
10186016	71	02/05/2014 00:00:00	12/08/2014 13:55:31	3.37 mths
10186644	55	01/05/2014 00:00:00	26/06/2014 00:00:00	1.84 mths

Case variant ID	Activity instances	Cases	Min duration	Median duration	Average duration	Max duration
1161	10	1	7 days	7 days	7 days	7 days
1162	9	1	3 days	3 days	3 days	3 days
1163	9	1	7.64 mths	7.64 mths	7.64 mths	7.64 mths
1164	9	1	1.17 mths	1.17 mths	1.17 mths	1.17 mths
1165	9	1	1.83 mths	1.83 mths	1.83 mths	1.83 mths
1166	7	1	1.78 wks	1.78 wks	1.78 wks	1.78 wks
1167	7	1	2.22 wks	2.22 wks	2.22 wks	2.22 wks
1168	6	1	3.47 days	3.47 days	3.47 days	3.47 days
1169	5	1	1.52 days	1.52 days	1.52 days	1.52 days
1170	3	1	3.51 days	3.51 days	3.51 days	3.51 days

Figura 3.38: Apromore - Más estadísticas del proceso obtenido como resultado

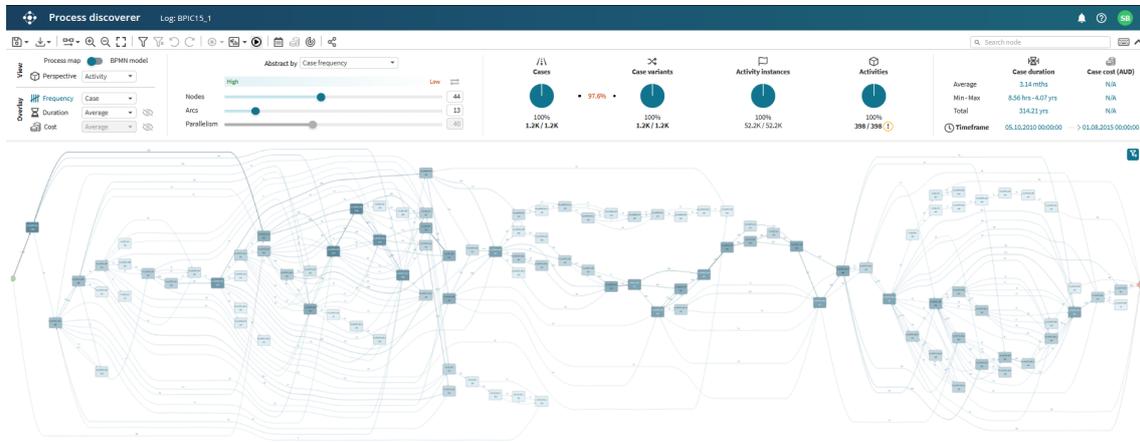


Figura 3.39: Apromore - Proceso representado gráficamente

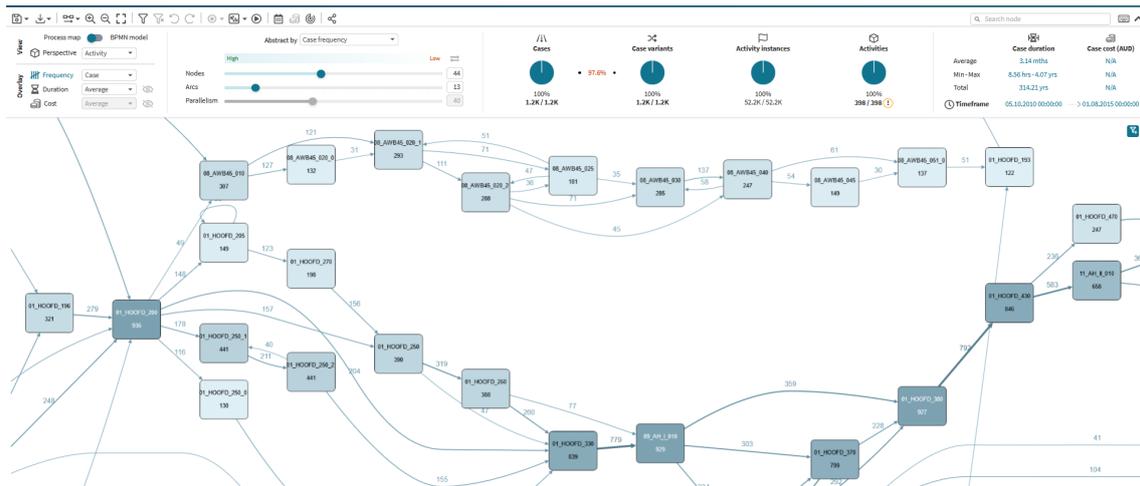


Figura 3.40: Apromore - Otra manera de representar el proceso gráficamente

CAPÍTULO 4

Relación con la carrera

Este trabajo se relaciona con la carrera de Ingeniería Informática en un alto grado, ya que, a lo largo del desarrollo del mismo, se utilizaron diversas tecnologías y entornos, que a lo largo de los 4 años de carrera también estuvieron presentes. Estas herramientas se comentan a continuación.

En primer lugar, se utilizaron entornos como github, el cuál es un repositorio online gratuito, que permite gestionar proyectos y controlar versiones de código, y es utilizado principalmente para almacenar trabajos, teniendo la posibilidad de compartir dicho trabajo con otras personas. Ahora para culminar con este entorno, hace falta mencionar que el control de versiones se refiere a que los usuarios/desarrolladores puedan ser capaces de administrar cambios en un software al mismo tiempo que el proyecto evoluciona, y para evitar realizar cambios sobre el código original, este control de versiones permite duplicar una parte del proyecto de forma aislada y da la oportunidad de trabajar sobre la misma, evitando así modificar el repositorio original [22].

Por otro lado, también se hizo uso del entorno .NET, que es una plataforma de código abierto para crear aplicaciones de escritorio, web y móviles que se puedan ejecutar de forma nativa en cualquier sistema operativo. Además, este entorno posee herramientas, bibliotecas y lenguajes que admiten el desarrollo de software moderno, entre los que destaca C Sharp, lenguaje de programación utilizado para llevar a cabo este trabajo, el cuál es orientado a objetos, y está relacionado con otros como C++, ya que C Sharp hereda muchas características de C++, pero con mejoras significativas [23].

Cabe destacar que, tanto github como el entorno .NET y C Sharp, son entornos y lenguajes de programación que se enseñaron en la materia cursada en tercer año de carrera denominada 'Ingeniería del Software', la cual me dio los conocimientos necesarios para que, a la hora de empezar a utilizarlas al inicio del desarrollo de este trabajo, supiera con qué iba a tratar y cómo manejar estos entornos y lenguajes de programación.

Como siguiente punto, también se utilizaron diversas aplicaciones basadas en minería de procesos, las cuáles son aplicaciones cuyo principal objetivo es el de analizar datos y mejorar los procesos. Esto relativo a minería de procesos también lo pude cursar en la materia denominada 'Sistemas de Información Estratégicos', materia correspondiente a la rama de 'Sistemas de Información', donde no entré en profundidad con el tema relativo a la minería de procesos, pero si que pude

utilizar aplicaciones de análisis de datos, como lo son PowerBI y RapidMiner5 [6].

Como último asunto respecto a este capítulo, en lo relativo a el tema de las competencias transversales, se debe destacar que, durante la elaboración de este TFG, las competencias transversales relativas al 'Trabajo en Equipo y Liderazgo' y 'Comunicación Efectiva' fueron requeridas y puestas en práctica en alto grado, ya que para poder llevar a cabo este trabajo, fue necesario trabajar en equipo en todo momento y el ser un líder, mediante diversas reuniones realizadas, donde se debían exponer cuestiones respecto a el cómo se iba a llevar a cabo el proyecto, que fases iba a tener, que se iba a desarrollar en cada una de ellas, que herramientas y tecnologías se iban a utilizar, entro otras cuestiones.

CAPÍTULO 5

Conclusiones y Trabajo Futuro

En el trabajo desarrollado se ha abordado el tema relativo a mejorar la calidad de los datos y la eficacia de los procesos en el ámbito sanitario, ámbito en donde la demanda de los servicios está en aumento constante, debido principalmente a que cada vez hay más gente mayor, lo que también genera un incremento de enfermedades crónicas. Esto comentado produce una presión considerable sobre los sistemas de salud, ya que se enfrentan a un nuevo desafío, en el cuál necesitan optimizar los recursos que poseen los mismos, con el fin de poder mantenerse a largo plazo. Teniendo en cuenta esto, se han tomado en cuenta metodologías como Lean Six Sigma y la minería de procesos para conseguir mejorar la calidad de los datos y la eficacia de los procesos.

En la propuesta de este trabajo, el estándar XES ha sido elemental, estándar del cuál se ha implementado un módulo para simplificar lo relativo al análisis de datos de eventos dentro de sistemas de información clínica. XES posibilita la detallada y estructurada documentación de los sucesos relacionados con la atención de los pacientes, lo cual es fundamental para garantizar la recopilación de la información requerida para llevar a cabo análisis eficaces. Esta aproximación posibilita que se entienda de mejor manera todo lo relativo a los flujos de trabajo y procesos clínicos, pudiendo ser capaces de identificar ineficiencias en estos procesos, como cuellos de botella y áreas de mejora.

El análisis comparativo realizado en este trabajo incluyó aplicaciones basadas en minería de procesos, como los son PMAApp, Celonis, ProM, Disco y Apromore, donde cada una de las mismas presenta sus puntos a favor y sus limitaciones, en cuestiones como la flexibilidad, facilidad de uso, tamaños de carga máximos, seguridad, entre otros aspectos. Por ejemplo, se puede decir que PMAApp, debido a su gran flexibilidad y el no tener límites de tamaño de archivo, es la app indicada para trabajar con grandes volúmenes de datos en entornos clínicos. Por otro lado, también cabe mencionar que Celonis ofrece herramientas avanzadas de análisis predictivo y monitoreo en tiempo real, aunque tenga limitaciones en cuanto a tamaños máximos de archivo y no sea para su uso exclusivo de usuarios pertenecientes al ámbito sanitario.

Sobre la metodología Lean Six Sigma, podemos añadir que, junto con aplicaciones basadas en minería de procesos, podría tener un impacto relevante en lo relativo al proceso de mejora de los procesos clínicos. No obstante, siempre hay que tener en cuenta que el factor decisivo es la calidad de los datos, ya que

actualmente, la mayoría de los problemas relativos al análisis de procesos clínicos se deben principalmente a datos incompletos o con poca exactitud. Por esto mencionado, el mejorar la calidad de éstos debe ser la principal prioridad para poder maximizar los beneficios que se puedan obtener de metodologías como la mencionada.

Con todo lo anterior, se sugiere seguir implementando sistemas que aseguren la calidad de los datos recopilados, con el fin de maximizar los beneficios del análisis de procesos clínicos, mediante minería de procesos y estándares como XES. También es muy importante el saber apostar por la capacitación del personal clínico y técnico, garantizando su comprensión sobre la relevancia de los datos de calidad y el empleo de herramientas avanzadas en su evaluación.

Como siguiente punto, se debe destacar que la combinación de estándares como XES con herramientas de minería de procesos ofrece un método efectivo para poder llevar a cabo la mejora de los procesos clínicos. No obstante, el éxito de esta combinación depende, principalmente, de la calidad de los datos y la formación del personal con respecto al uso de estas tecnologías. Este estudio proporciona el fundamento para investigaciones y avances futuros en este ámbito, con el principal objetivo de mejorar la atención al paciente y la sostenibilidad de los sistemas de salud.

Para culminar con este capítulo, cabe señalar que este proyecto se llevó a cabo en el grupo de investigación SABIEN (Innovaciones Tecnológicas para la Salud y el Bienestar) del Instituto Universitario de Tecnologías de la Información y Comunicaciones (ITACA), centro de excelencia en la investigación y desarrollo de las TIC, con vocación de innovación y de transferencia de sus productos y servicios de I+D+i a los sectores públicos, industrial, comercio y terciario. El instituto ITACA, adscrito a la Universitat Politècnica de València (UPV), está ubicado en la Ciudad Politécnica de la Innovación (CPI). En este caso, junto a este grupo mencionado, se implementaron los dos módulos del estándar XES para la aplicación PMApp, como también se llevó a cabo una comparativa realizada entre aplicaciones de minería de procesos, y en un futuro se pretende seguir colaborando con ellos, para la realización de próximos trabajos futuros [24].

Bibliografía

- [1] World Health Organization and others, *Ageing*, OECD, 2020. Accedido: 27 de agosto de 2024.
- [2] Pepper, Matthew PJ and Spedding, Trevor A, "The evolution of lean Six Sigma", *International Journal of Quality & Reliability Management*, Emerald Group Publishing Limited, 2010. Accedido: 27 de agosto de 2024.
- [3] Hayes, Sherrill H and Carroll, Susan R, "Early intervention care in the acute stroke patient", *Archives of physical medicine and rehabilitation*, vol. 67, no. 5, pp. 319–321, Elsevier, 1986.
- [4] King, Rosemarie B, "Quality of life after stroke", *Stroke*, vol. 27, no. 9, pp. 1467–1472, Am Heart Assoc, 1996.
- [5] Martin, Niels, "Data quality in process mining", in *Interactive Process Mining in Healthcare*, Springer, 2021, pp. 53–79.
- [6] Van Der Aalst, W., "Process mining," *Communications of the ACM*, vol. 55, no. 8, pp. 76–83, 2012.
- [7] Acampora, Giovanni, Vitiello, Autilia, Di Stefano, Bruno, van der Aalst, Wil and Günther, Christian, "The xes standard", *IEEE Computational Intelligence Magazine*, pp. 4–8, 2017.
- [8] Celonis, *Celonis: Software de Minería de Procesos*. Disponible en: <https://www.celonis.com/es/>, Accedido: 23 de septiembre de 2024.
- [9] Van der Aalst, Wil MP, van Dongen, Boudewijn F, Günther, Christian W, Rozinat, Anne, Verbeek, Eric and Weijters, Ton, "ProM: The process mining toolkit.", *BPM (Demos)*, vol. 489, no. 31, pp. 2, 2009.
- [10] Fluxicon, *Disco: Process Mining Software*. Disponible en: <https://fluxicon.com/disco/>, Accedido: 23 de septiembre de 2024.
- [11] Apromore, *Apromore: Key Features of Process Mining Software*. Disponible en: <https://apromore.com/key-features>, Accedido: 23 de septiembre de 2024.
- [12] Ibanez-Sanchez, Gema, Carlos Fernandez-Llatas, Zoe Valero-Ramon, y Jose Luis Bayo-Monton. *Pmapp: an interactive process mining toolkit for building healthcare dashboards*. En *International Workshop on Explainable Artificial Intelligence in Healthcare*, págs. 75–86, 2023. Springer.

- [13] Corrigan, Janet M, et al. *Crossing the Quality Chasm*. Building a Better Delivery System, vol. 89, 2005.
- [14] Simal Hernández, Patricia, Jorge Matías Guiu-Guia, Teresa Hernández Meléndez, y Pilar Aparicio Azcárraga. *Logros y retos en la atención del ictus en España: desde la estrategia del sistema nacional de salud al plan de acción europeo 2018-2030*. Revista Española de Salud Pública, vol. 95, pp. perspectivas 21, 2022, SciELO Public Health.
- [15] Mackway-Jones, Kevin, Janet Marsden, y Jill Windle. *Emergency Triage: Manchester Triage Group*. John Wiley & Sons, 2013.
- [16] Ministerio de Sanidad, Consumo y Bienestar Social, *Informe Anual del Sistema Nacional de Salud 2019*, MINISTERIO DE SANIDAD, 2021. Disponible en: <https://www.mscbs.gob.es/estadEstudios/estadisticas/sisInfSanSNS/tablasEstadisticas/InfAnSNS.htm>. Accedido: 23 de septiembre de 2024
- [17] De Koning, Henk, John PS Verver, Jaap van den Heuvel, Soren Bisgaard, y Ronald JMM Does. *Lean Six Sigma in Healthcare*. Journal for Healthcare Quality, vol. 28, no. 2, pp. 4–11, 2006, Wiley Online Library.
- [18] Snee, Ronald D. *Lean Six Sigma—Getting Better All the Time*. International Journal of Lean Six Sigma, 2010, Emerald Group Publishing Limited.
- [19] Dash, Sabyasachi, Sushil Kumar Shakyawar, Mohit Sharma, y Sandeep Kaushik. *Big Data in Healthcare: Management, Analysis and Future Prospects*. Journal of Big Data, vol. 6, no. 1, pp. 1–25, 2019, Springer.
- [20] Kim, Christopher S, Spahlinger, David A, Kin, Jeanne M and Billi, John E, "Lean health care: what can hospitals learn from a world-class automaker?", *Journal of Hospital Medicine: an official publication of the Society of Hospital Medicine*, vol. 1, no. 3, pp. 191–199, Wiley Online Library, 2006.
- [21] Toussaint, John S and Berry, Leonard L, "The promise of lean in health care", in *Mayo clinic proceedings*, vol. 88, no. 1, pp. 74–82, Elsevier, 2013.
- [22] Blischak, John D, Davenport, Emily R and Wilson, Greg, "A quick introduction to version control with Git and GitHub", *PLoS computational biology*, vol. 12, no. 1, pp. e1004668, Public Library of Science San Francisco, CA USA, 2016.
- [23] Nagel, Christian, *Professional C# and .Net*, John Wiley & Sons, 2021.
- [24] SABIEN, *¿Qué es SABIEN?*. Disponible en: <http://www.sabien.upv.es/presentacion/que-es-sabien/>, Accedido: 23 de septiembre de 2024.
- [25] Vincent, Charles, *Patient Safety*. John Wiley & Sons, 2011.
- [26] XES Standard - eXtensible Event Stream, <https://www.xes-standard.org/>, Accedido: 25 de septiembre de 2024.

APÉNDICE A

ODS



ANEXO

OBJETIVOS DE DESARROLLO SOSTENIBLE

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No Procede
ODS 1. Fin de la pobreza.				X
ODS 2. Hambre cero.				X
ODS 3. Salud y bienestar.	X			
ODS 4. Educación de calidad.				X
ODS 5. Igualdad de género.				X
ODS 6. Agua limpia y saneamiento.				X
ODS 7. Energía asequible y no contaminante.				X
ODS 8. Trabajo decente y crecimiento económico.				X
ODS 9. Industria, innovación e infraestructuras.			X	
ODS 10. Reducción de las desigualdades.				X
ODS 11. Ciudades y comunidades sostenibles.			X	
ODS 12. Producción y consumo responsables.				X
ODS 13. Acción por el clima.				X
ODS 14. Vida submarina.				X
ODS 15. Vida de ecosistemas terrestres.				X
ODS 16. Paz, justicia e instituciones sólidas.				X
ODS 17. Alianzas para lograr objetivos.				X

T +34 963 877 210 F +34 963 877 219





Reflexión sobre la relación del TFG/TFM con los ODS y con el/los ODS más relacionados.

Este trabajo presente nace de la necesidad de mejorar la calidad de los datos clínicos y de la mejora de los distintos procesos relativos al mismo ámbito, teniendo como principal objetivo el de hacer que los sistemas de salud puedan permanecer con el tiempo, aunque la demanda de éstos cada vez sea mayor.

Una vez comentado esto, se puede considerar que el presente proyecto está altamente relacionado con la ODS número tres (3), la cual consiste en “Salud y Bienestar”, ya que el módulo implementado del estándar ya detallado en este documento está estrictamente ligado al tema relativo a mejorar la calidad de los datos clínicos, ya que dicho módulo, junto con la ayuda de la aplicación PMAApp ya comentada, tiene como función principal el analizar los datos que se le pasen en forma de archivo, y proporcionar una salida donde se pueda visualizar todo lo relativo al archivo, incluyendo hasta los mínimos detalles, evitando así la pérdida de información relevante y pudiendo contribuir a mejorar la calidad de los datos, ayudando también a disminuir posibles colapsos que se puedan generar en los distintos sistemas de salud, lo que es particularmente relevante para esta ODS, ya que los colapsos en los sistemas de salud pueden tener consecuencias devastadoras para la salud pública, y por lo tanto, al asegurar que los datos sean precisos y fácilmente accesibles, este proyecto ayuda a los profesionales de la salud a gestionar los recursos más eficazmente, lo que puede prevenir situaciones de crisis, las cuáles se quieren evitar a toda costa en este tipo de sistemas tan críticos. Por lo tanto, y para concluir con esta primera ODS, teniendo datos clínicos precisos y bien gestionados es fundamental para responder a emergencias de salud pública de manera eficaz.

T +34 963 877 210 F +34 963 877 219



Por otro lado, este proyecto también se puede relacionar mínimamente con la ODS número nueve (9), la cual consiste en “Industria, Innovación e Infraestructura ” ya que, además de implementar el módulo ya mencionado, también se utilizan las últimas tecnologías, en este caso, aplicaciones de minería de procesos, y, sobre todo, la denominada PMAApp, para contribuir a esta mejora en la calidad de los datos y en la mejora de los procesos clínicos. Este uso de este tipo de aplicaciones también contribuye a la creación de una infraestructura digital más robusta y resiliente, capaz de soportar grandes volúmenes de datos clínicos y mejorar la gestión de la información en los sistemas de salud, lo que hoy en día es fundamental en este tipo de sistemas.

Por último, la ODS número once (11) que consiste en “Ciudades y Comunidades sostenibles” también se relaciona levemente con este proyecto, ya que el mismo, al estar centrado en mejorar la calidad de los datos clínicos y optimizar los procesos de este ámbito, puede tener un impacto significativo en la calidad de los servicios de salud que se ofrecen a las comunidades, ya que las ciudades y comunidades sostenibles requieren de servicios de salud eficientes y accesibles que sean capaces de dar respuesta de manera rápida a las necesidades poblacionales.

T +34 963 877 210 F +34 963 877 219

