

Documento de Arquitectura de Software

SleepSmart: Prototipo para el pre diagnóstico de trastornos del sueño a partir de datos de salud.

Elaborado por:
Felix Rafael Moreno Tabares
Juliana Murillo Marmolejo

Pontificia Universidad Javeriana
Bogotá, 2023

Historial

Fecha	Versión	Descripción
28-09-2023	1.0	Iteración inicial
18-10-2023	1.1	Reestructuración del documento

Tabla de contenido

1. Introducción.....	5
1.1. Propósito	5
1.2. Alcance	6
1.3. Definiciones, Acrónimos y Abreviaciones	6
1.4. Referencias	6
1.5. Descripción General.....	6
2. Objetivos Arquitectónicos y Restricciones	7
2.1. Plataforma Técnica	7
2.2. Restricciones	8
3. Representación Arquitectónica.....	8
3.1. Requerimientos Arquitectónicamente Significativos	9
4. Diseño	10
4.1. Justificación.....	10
4.2. Vista Lógica	12
4.3. Vista Implementación.....	13
4.4. Vista Proceso.....	15
4.5. Vista Despliegue	19
5. Tamaño y Rendimiento	20
6. Calidad	20

Índice de Ilustraciones

Ilustración 1 - Modelo Vistas 4+1	5
Ilustración 2 - Vista Lógica	12
Ilustración 3 - Diagrama de Componentes	14
Ilustración 4 - Diagrama Secuencia Entrenamiento	15
Ilustración 5 - Diagrama Secuencia Registro Usuario.....	16
Ilustración 6 - Diagrama Secuencia Cargar Datos Médicos	16
Ilustración 7 - Diagrama Secuencia Análisis de Datos Médicos.....	17
Ilustración 8 - Diagrama Secuencia Reporte Análisis Resumen	17
Ilustración 9 - Diagrama Secuencia Reporte Análisis Medico Completo	18
Ilustración 10 - Diagrama Secuencia Reporte Métricas	18

Índice de Tablas

Tabla 1 - Comparativa Patrones Arquitecturales.....	11
---	----

1. Introducción

Este documento surge como un complemento al documento de especificación de requerimientos y se busca ilustrar la estructura del proyecto para vigilancia y seguimiento de historias clínicas en pacientes durante su fase de diseño arquitectónico.

Actividades arquitecturales	SAD
Paso 1 – Identificar los objetivos y restricciones arquitecturales	Sección 2
Paso 2 – Definir arquitecturas candidatas	Sección 3
Paso 3 – Diseñar la arquitectura objetivo	Sección 4
Paso 4 – Documentar las diferentes vistas	Sección 4
Paso 5 – Identificar los datos	Sección 5

1.1. Propósito

El documento de especificación de la arquitectura (SAD) provee una perspectiva general del sistema Sleep Monitor. Este presenta variadas vistas arquitecturales para representar diferentes aspectos del sistema. La intención es capturar y conciliar las diferentes decisiones arquitectónicas que se han tomado para diseñar el sistema. Para representar el sistema Sleep Monitor de forma más precisa, la estructura de este documento estará basada en el modelo “4+1” de Kruchten.



Ilustración 1 - Modelo Vistas 4+1

Este modelo de vistas “4+1” permitirá a diferentes tipos de Stakeholders encontrar lo que necesiten dentro del diseño arquitectónico.

1.2. Alcance

El prototipo SleepSmart, se encargará del procesamiento de reportes médicos generados por los relojes inteligentes relacionados con el sueño, en el cual estos serán almacenados en un lago de datos, el cual estarán ubicados en la nube privada y enviados al componente analítico para su procesamiento en el modelo Light Gradient Boosting Machine (LightGBM); posterior almacenado el resultado para la generación de métricas reportes según el usuario consultante.

Este sistema podría usarse para generar prediagnósticos con los datos médicos de pacientes en tiempo real o con post procesamiento:

- Clasificación de pacientes.
- Monitoreo de histórico

1.3. Definiciones, Acrónimos y Abreviaciones

- UML: Lenguaje de modelado unificado
- SAD: Documento de arquitectura de software

1.4. Referencias

- The “4+1” view model of software architecture, Philippe Kruchten, November 1995

1.5. Descripción General

La relación entre el paciente y el médico en el contexto del análisis del sueño a través de datos capturados por relojes inteligentes se ha transformado en una colaboración más accesible y conveniente. Estos dispositivos permiten que los pacientes monitoreen su sueño de manera continua y sin intervenciones invasivas, lo que reduce la necesidad de evaluaciones especializadas que pueden resultar molestas. Los datos recopilados por los relojes inteligentes proporcionan información valiosa sobre los patrones de sueño y los síntomas asociados a posibles trastornos, lo que facilita la comunicación con los médicos.

Esta nueva dinámica de paciente y médico se basa en la recopilación de datos objetivos y en la posibilidad de compartir fácilmente la información con el profesional de la salud. Los pacientes pueden tener un papel más activo en su propio cuidado al registrar y compartir los datos de su sueño, lo que a su vez permite a los médicos realizar evaluaciones más informadas y personalizadas. Esto puede mejorar la calidad de la atención y facilitar el diagnóstico y tratamiento temprano de los trastornos del sueño. Además, esta tecnología puede contribuir a una atención más eficiente y asequible en el sistema de salud.

2. Objetivos Arquitectónicos y Restricciones

Esta sección describe los objetivos y requerimientos que tienen un impacto significativo en la arquitectura.

2.1. Plataforma Técnica

La elección de la plataforma técnica es un aspecto crítico en el diseño y desarrollo del prototipo. La plataforma técnica se refiere a las tecnologías, herramientas y componentes que se utilizarán para construir el sistema.

- **Base de Datos:** La selección de la base de datos es un elemento crítico para el almacenamiento y recuperación de datos en el sistema. Debe realizarse cuidadosamente teniendo en cuenta las características de los datos y los requisitos específicos del proyecto. En el caso del prototipo, se ha optado por utilizar dos bases de datos, cada una con un enfoque específico; la primera base de datos se centra en la gestión de datos estructurados, como la información de los usuarios. En este contexto, se han considerado opciones relacionales y la segunda base de datos se utiliza para el manejo de datos generados por los relojes inteligentes relacionados con el sueño. Estos datos suelen ser no estructurados o semiestructurados, lo que hace que las bases de datos NoSQL.
- **Lenguaje de Programación:** La selección del lenguaje de programación es crítica en el diseño, y es esencial considerar cuidadosamente las características y ventajas de cada lenguaje candidato. El prototipo requiere un lenguaje robusto en análisis de datos y flexibilidad en el Backend para lograr un procesamiento eficaz de la información generada por los relojes inteligentes y, al mismo tiempo, garantizar la escalabilidad y mantenimiento del sistema.
- **Frameworks:** Desempeñan un papel fundamental en el proceso de desarrollo al permitir una programación más eficiente y el cumplimiento de buenas prácticas. En el ámbito del análisis de datos, es posible considerar la adopción de bibliotecas especializadas que faciliten la implementación de soluciones de aprendizaje automático o la gestión de datos.
- **Seguridad:** Se deben seleccionar herramientas y prácticas de seguridad para proteger los datos de los usuarios y garantizar la integridad y confidencialidad de la información médica. Esto puede incluir la implementación de firewalls, encriptación de datos y autenticación.
- **Cloud:** Por motivos inherentes a la naturaleza del prototipo, no se considera recomendable utilizar una nube privada para el almacenamiento de los datos médicos en el Datalake. En su lugar, se debe evaluar la utilización de una solución

que garantice la seguridad, privacidad y cumplimiento normativo de estos datos sensibles, lo que podría incluir la adopción de prácticas de seguridad adicionales o la exploración de alternativas de almacenamiento que cumplan con los estándares médicos y regulaciones de privacidad aplicables.

- **Hardware:** La especificación de requisitos de hardware es un paso fundamental para garantizar que el sistema pueda satisfacer sus necesidades tanto en la nube privada como en la pública. Esto implica definir con precisión los recursos requeridos, como servidores con capacidades de procesamiento adecuadas, capacidad de almacenamiento escalable para acomodar el crecimiento de datos, y recursos de memoria y red para garantizar un rendimiento óptimo.
- **Escalabilidad:** La plataforma técnica debe ser capaz de escalar horizontal o verticalmente para manejar un aumento en la carga de trabajo.

2.2. Restricciones

La restricción más significativa y crítica que impone un fuerte impacto en el diseño del sistema es la decisión de no utilizar una nube privada para el manejo de datos médicos en el Datalake. Esta restricción es un resultado directo de las estrictas implicaciones de seguridad y privacidad que están inherentemente vinculadas a los datos médicos. La elección de evitar el uso de una nube privada se motiva por la necesidad apremiante de cumplir con regulaciones altamente específicas de privacidad de datos médicos, como la normativa HIPAA en los Estados Unidos o el habeas data en Colombia.

3. Representación Arquitectónica

Este documento detalla la arquitectura del sistema utilizando las vistas definidas en el modelo "4+1". Las vistas utilizadas para documentar las aplicaciones relacionadas con el sistema Sleep Monitor son:

Vista Lógica

- Audiencia: Diseñadores
- Área: Requerimientos funcionales: describe los casos de uso más significativos.
- Artefactos relacionados: Modelo de diseño Philippe Kruchten, November 1995

Vista de procesos

- Audiencia: Diseñadores
- Área: Requerimientos no funcionales: describe los aspectos de concurrencia y sincronización en el modelo "4+1". Las vistas utilizadas para documentar las aplicaciones relacionadas con la aplicación.
- Artefactos relacionados: (sin artefacto específico)

Vista de implementación

- Audiencia: Desarrolladores funcionales: describe los casos de uso más significativos.
- Área: Componentes de software: se describen las capas y subsistemas de la aplicación
- Artefactos relacionados: Modelo de implementación y componentes

Vista de despliegue

- Audiencia: Desarrolladores funcionales: Describe los aspectos de concurrencia y sincronización
- Área: Topología; describe el mapeo del software en el hardware y muestra los aspectos distribuidos del sistema.
- Artefactos relacionados: Modelo de despliegue

Vista de casos de uso

- Audiencia: Todos los Stakeholders del sistema, incluyendo usuarios finales de la aplicación
- Área: Describe el conjunto de escenarios que representan una funcionalidad significativa o central del sistema
- Artefactos relacionados: Modelo de casos de uso, documentos de casos de usos

Vista de datos

- Audiencia: Desarrolladores funcionales, analistas de datos en el hardware y muestra los aspectos distribuidos del sistema.
- Área: Persistencia: describe los elementos de persistencia significativos en el modelo de datos
- Artefactos relacionados: Modelo de datos

3.1. Requerimientos Arquitectónicamente Significativos

Cumplimiento de las Restricciones

- El acceso a la información de cada paciente debe ser restringido por el control de acceso siendo disponible solo para el paciente y personal médico que lo esté tratando.
- Toda la información de aspecto científico, técnico y administrativo relativos a la atención en salud debe estar compilada en la historia clínica.
- Se debe permitir acceso a la edición inmediata o simultánea al prestar el servicio.
- La historia clínica debe ser secuencial y marcada con números para conocer la secuencia en la que fue realizada.

Fiabilidad

- Precisión en el modelo de inferencia del 90%
- La aplicación debe estar disponible el 99,9% del tiempo del año disponible.
- La aplicación debe ser capaz de soportar caídas en los servidores y continuar operando con normalidad.

Seguridad

- Los datos para transportar deben estar encriptados para que no puedan ser vistos más que por las partes autorizadas.

Mantenibilidad

- Se deben poder agregar nuevas estrategias de entrenamiento en la generación de modelos de inferencias.

Compatibilidad

- Se refiere a la compatibilidad del sistema para poder recibir datos exportados de otros relojes inteligentes. Esto implica que el sistema debe estar diseñado y configurado de manera que pueda recibir y procesar datos provenientes de diferentes marcas o modelos de relojes inteligentes, en lugar de estar restringido a uno o algunos modelos específicos.

4. Diseño

4.1. Justificación

Para satisfacer los requerimientos de la arquitectura se tomaron diferentes estilos arquitectónicos mezclados en diferentes partes del sistema. Para manejar la interactividad de los clientes por medio de browser se planeó un modelo MVC. En cuanto a las funcionalidades de manejo datos, registro y la analítica se plantea un estilo por capas en el que la primera capa es la fachada de servicios, estos acceden a módulos que procesan la información y es guardada en la base de datos y es doble vía. Allí se pierde rendimiento, pero se gana fiabilidad y se implementa la seguridad como una capa transversal además de módulos para el monitoreo y manejo de excepciones que contribuyen a la disponibilidad y seguridad necesarios en la aplicación. La restricción principal está dada por la naturaleza de los datos; por ende, la aplicación estará en una nube privada.

El desarrollo consta de dos aplicaciones fundamentales: la aplicación de Backoffice y la aplicación de Analítica/Entrenamiento. La aplicación de Negocio se encarga de la recepción de datos generados los relojes inteligentes, así como de la interacción con los usuarios finales por medio de la interfaz web, asegurando una experiencia fluida y

efectiva. Por otro lado, la aplicación de Analítica/Entrenamiento se especializa en el procesamiento y análisis de datos, así como en el entrenamiento de algoritmos de aprendizaje automático para la clasificación de estos. Esta estructura de dos aplicaciones permite una clara separación de responsabilidades y una colaboración eficiente entre la adquisición de datos y la generación de conocimiento en el análisis de sueño a partir de datos de relojes inteligentes.

A continuación, se detallan en la siguiente tabla 1 la selección de los patrones arquitectónicos que pueden ser candidatos a usar dentro del desarrollo de la aplicación, se utilizó la priorización de los atributos de calidad y con esta base, se procedió a seleccionar los atributos para los diferentes subsistemas que se podían llegar a ser utilizados y a su vez favorecen los atributos de calidad más importantes del sistema.

*Tabla 1 - Comparativa Patrones Arquitecturales
 Elaboración Propia*

Patrones	Atributos Positivos	Atributos Negativos	Pacientes	Personal Salud	Organizaciones de Salud
SideCar	Modularidad Flexibilidad Escalabilidad				x
API Gateway	Usabilidad Mantenibilidad	Disponibilidad	x	x	x
Microservicios	Escalabilidad Mantenibilidad	Seguridad	x	x	x
Pipes and filter (protocol pipeline)	Flexibilidad Interoperabilidad Reusabilidad	Rendimiento Usabilidad		x	
Layers	Seguridad Mantenibilidad Fiabilidad	Compatibilidad	x	x	x

El uso del patrón de diseño de capas (Layers) en un prototipo de análisis de sueño con datos de relojes inteligentes ofrece una serie de beneficios cruciales para el desarrollo exitoso de esta aplicación. En primer lugar, la estructura de capas permite una clara separación de responsabilidades y funciones, lo que simplifica el desarrollo y mantenimiento.

La aplicación Backoffice en este contexto está organizada en cuatro capas distintas: Presentación, Negocio, Integración. La capa de Presentación proporciona la interfaz a través de la cual los usuarios interactúan con la aplicación. La capa de negocio constituye el núcleo de la aplicación, donde se encuentra la lógica y las reglas que procesan los datos de sueño para almacenarlos en el lago de datos y posteriormente enviarlo a la aplicación de analítica. La capa de Integración se encarga de gestionar la interacción con otros sistemas (Analítica) y bases de datos. Esta estructura de capas optimiza la organización y el rendimiento de la aplicación, lo que resulta fundamental en un prototipo de análisis de sueño a partir de datos de relojes inteligentes.

Por el lado de la analítica, la aplicación consta de tres capas fundamentales: Presentación, Negocio e Integración. La capa de Presentación está diseñada para permitir la interacción exclusiva de los expertos, facilitando el proceso de entrenamiento de los modelos de inferencia por medio de algoritmos predefinidos. Aquí, los expertos pueden cargar datos, configurar parámetros y supervisar el progreso del entrenamiento. La capa de Negocio se encarga de exponer los servicios y funcionalidades necesarios para llevar a cabo el análisis de los datos de sueño. Esta capa es esencial para aplicar algoritmos de aprendizaje automático y extraer información significativa de los datos recopilados a través de los relojes inteligentes. La capa de Integración juega un papel crucial al conectarse con servicios como MLflow, que facilitan la gestión y el seguimiento de los modelos de aprendizaje automático.

4.2. Vista Lógica

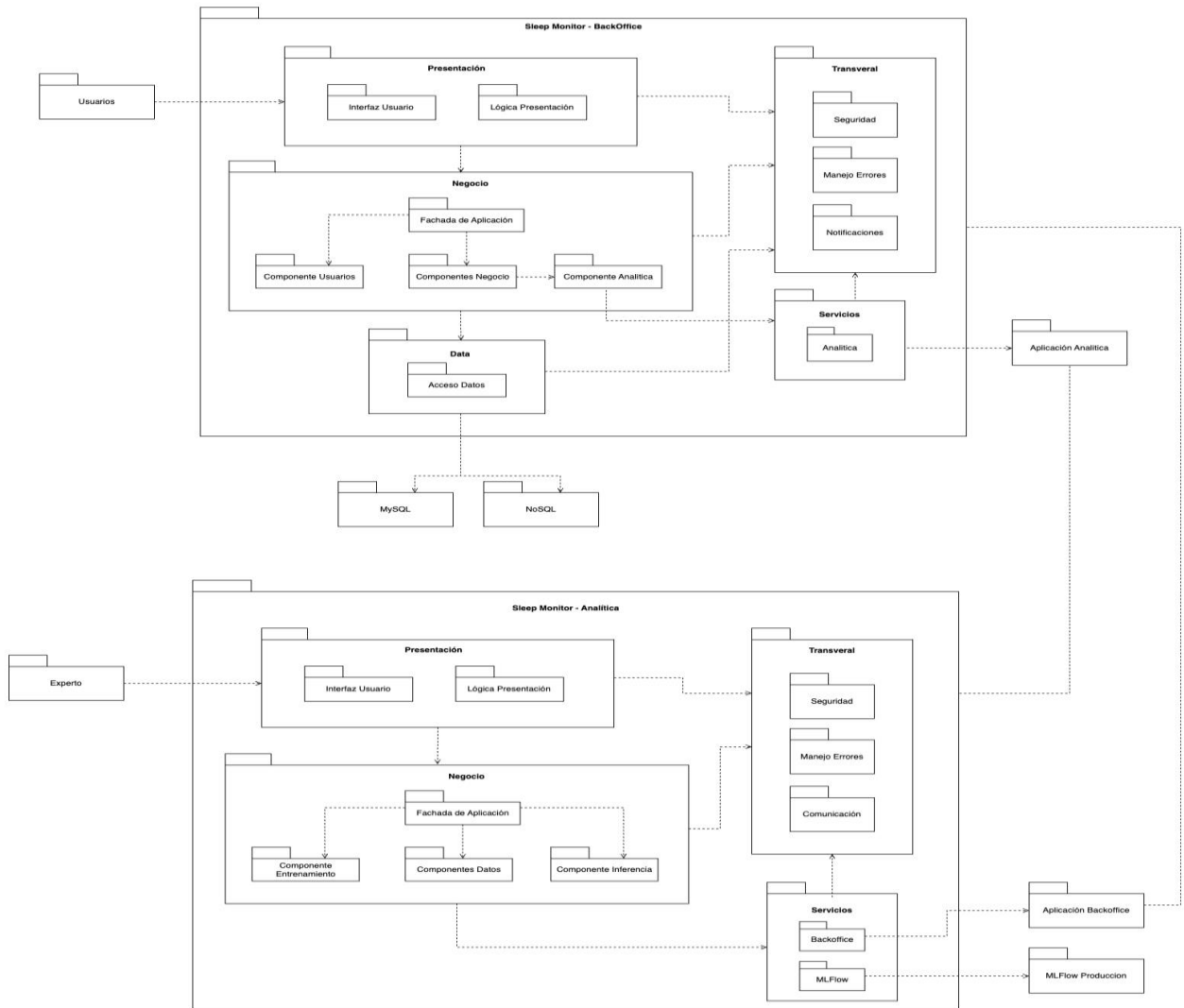


Ilustración 2 - Vista Lógica
Elaboración Propia

La vista lógica del prototipo Sleep Monitor se compone de dos aplicaciones interconectados que colaboran para ofrecer una solución integral de análisis de los datos enfocados en el insomnio. La primera aplicación denominada "Sleep Monitor BackOffice," tiene la responsabilidad de administrar el registro de usuarios, recibir y procesar los datos provenientes de los relojes inteligentes, y presentar de manera accesible los resultados del análisis realizado por el modelo de inferencia.

La segunda aplicación denominada "Sleep Monitor Analítica," desempeña un papel fundamental en la evolución y mejora continua del sistema. En esta sección, los usuarios pueden entrenar el modelo de inferencia utilizando el algoritmo LightGBM. Este proceso se basa en los datos almacenados en el lago de datos, disponible en el componente anterior. Una vez que el modelo ha sido entrenado, esta sección también habilita a los usuarios a realizar análisis de la inferencia, haciendo uso de los datos previamente cargados en el componente BackOffice. Esta funcionalidad proporciona una visión detallada de la calidad del sueño y permite a los usuarios tomar decisiones más fundamentadas para mejorar sus patrones de descanso. En conjunto, estos dos componentes forman un sistema completo y eficaz para el monitoreo y la gestión del sueño de los usuarios.

4.3. Vista Implementación

Esta vista detallada de cómo se materializa la arquitectura de software en términos de componentes, módulos y tecnologías específicas. Esta sección se adentra en la implementación concreta de la arquitectura definida, destacando las decisiones clave de diseño, la estructura del código, la selección de herramientas y tecnologías, y otros aspectos cruciales que transforman la visión arquitectónica en una realidad funcional. En esencia, la "Vista de Implementación" conecta la teoría arquitectónica con la práctica con el fin de guiar a los desarrolladores y demás partes interesadas en la construcción efectiva del software.

A continuación, se presenta la ilustración 3 el diagrama de componentes, donde se destacan tres elementos clave que son fundamentales del sistema Sleep Monitor: la "Aplicación Web", "BackOffice" y "Analítica".

La "Aplicación Web" desempeña un papel central al permitir a los usuarios interactuar con el sistema de manera efectiva. Esta aplicación se ha desarrollado utilizando Vite React, el cual es consiste en entorno de desarrollo ultrarrápido para aplicaciones web, y React es una biblioteca JavaScript ampliamente utilizada para creación de interfaces. Entre las principales funcionalidades de esta aplicación se incluyen el registro de usuarios, la capacidad de cargar datos para su posterior análisis, la generación de reportes detallados basados en el resultado de la inferencia de estos datos, así como la presentación de métricas relevantes a las organizaciones de salud. Su interfaz amigable y su accesibilidad a través de un navegador web la convierten en una herramienta clave para los usuarios que desean comprender su sueño de manera efectiva.

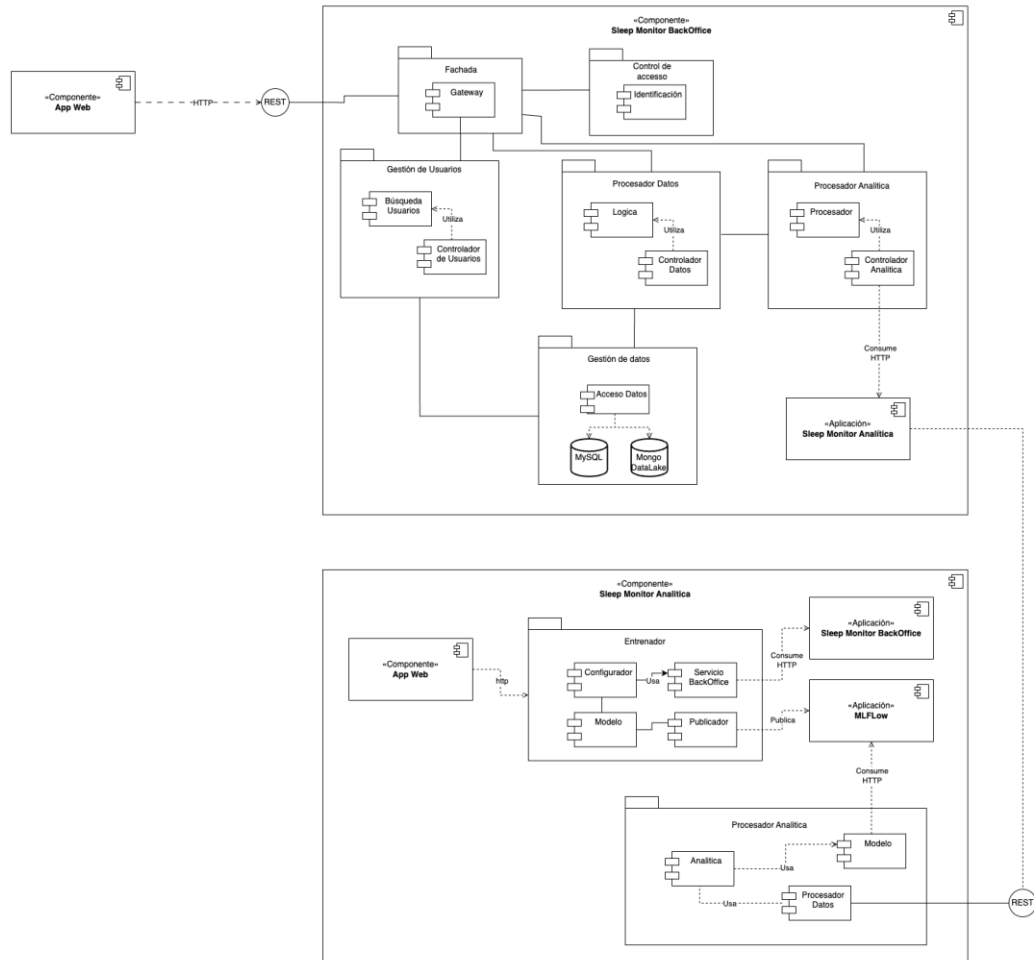


Ilustración 3 - Diagrama de Componentes
Elaboración Propia

Por otra parte, El "Sleep Monitor BackOffice" desempeña un papel crucial en la gestión y procesamiento de datos para el sistema. Está desarrollado utilizando el lenguaje de programación GO, conocido por su eficiencia y capacidad para manejar tareas relacionadas con el rendimiento y la concurrencia. Se cuenta con el componente de la gestión de usuarios, permitiendo el registro y administración de perfiles, incluyendo autenticación y asignación de permisos, utilizando una base de datos MySQL. Además, permite el cargue de los datos de relojes inteligentes, los procesa y los almacena en un sistema de tipo "Datalake" en MongoDB, donde quedan listos para su análisis posterior. Además, se conecta vía API con el componente de "Sleep Monitor Analítica" para permitir un análisis más profundo y enriquecedor, generando datos relevantes que proporcionan a los usuarios la información necesaria para comprender sus patrones de sueño y tomar decisiones bien fundamentadas en aras de mejorar su calidad de descanso.

Por último, el "Sleep Monitor Analítica" es el componente fundamental dentro del sistema, brindando a al experto de datos la capacidad de entrenar el modelo de inferencia mediante el uso del algoritmo LightGBM, aprovechando los datos almacenados en el Datalake del BackOffice. Además, facilita el análisis de la inferencia, empleando el modelo publicado en MLFLOW tras completar el proceso de entrenamiento. Esta funcionalidad permite que los usuarios obtengan recomendaciones personalizadas basadas en un análisis de los datos relacionados con sus patrones de sueño. La conexión directa entre el BackOffice y el componente de Analítica garantiza una sinergia efectiva que posibilita la optimización constante del sistema, contribuyendo a una experiencia más informada y satisfactoria para los usuarios de Sleep Monitor.

4.4. Vista Proceso

En esta vista, los diagramas de secuencia se convierten en una herramienta esencial para describir el flujo de información y control a medida que las entidades del sistema cooperan para lograr una funcionalidad específica. Estos diagramas permiten una comprensión detallada de los eventos y procesos que ocurren en el sistema y garantizar que el software cumpla con los requisitos del usuario de manera efectiva.

A continuación, se detallan siete (7) diagramas de procesos. Cada uno de estos diagramas visualiza el procesos específico, desde el entrenamiento del modelo de inferencia hasta la generación de informes y métricas. Estas representaciones visuales ofrecen una comprensión clara de cómo las distintas partes del sistema interactúan en tareas clave, como el registro de usuarios, el análisis de datos de sueño, y la presentación de resultados, contribuyendo a una implementación efectiva y a una experiencia de usuario informada y satisfactoria.

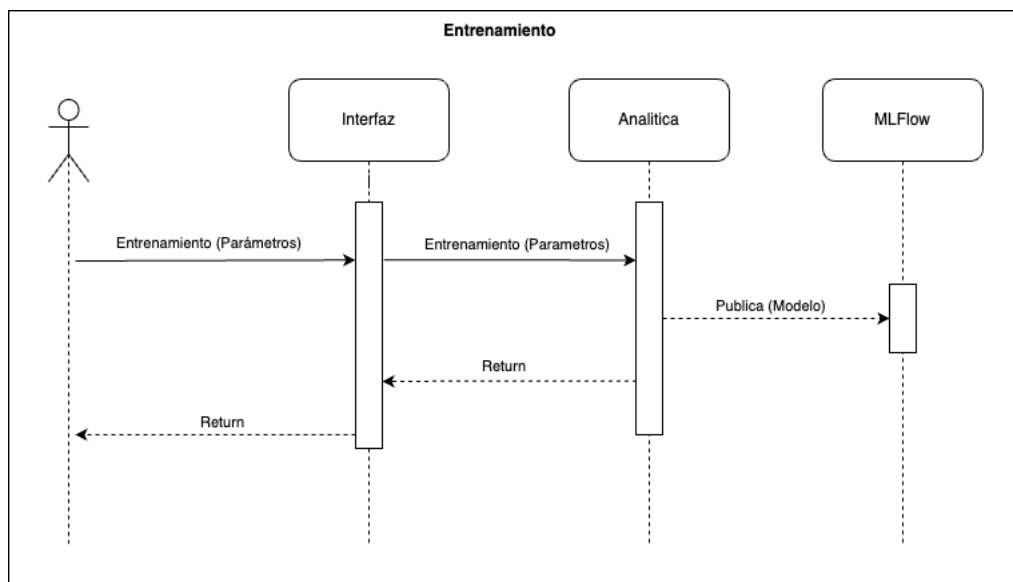


Ilustración 4 - Diagrama Secuencia Entrenamiento
Elaboración Propia

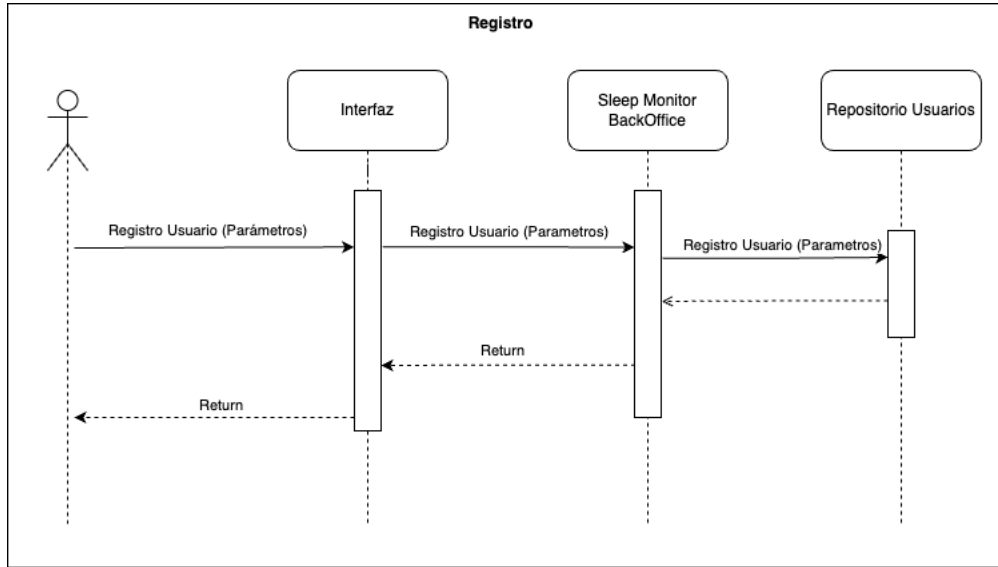


Ilustración 5 - Diagrama Secuencia Registro Usuario
Elaboración Propia

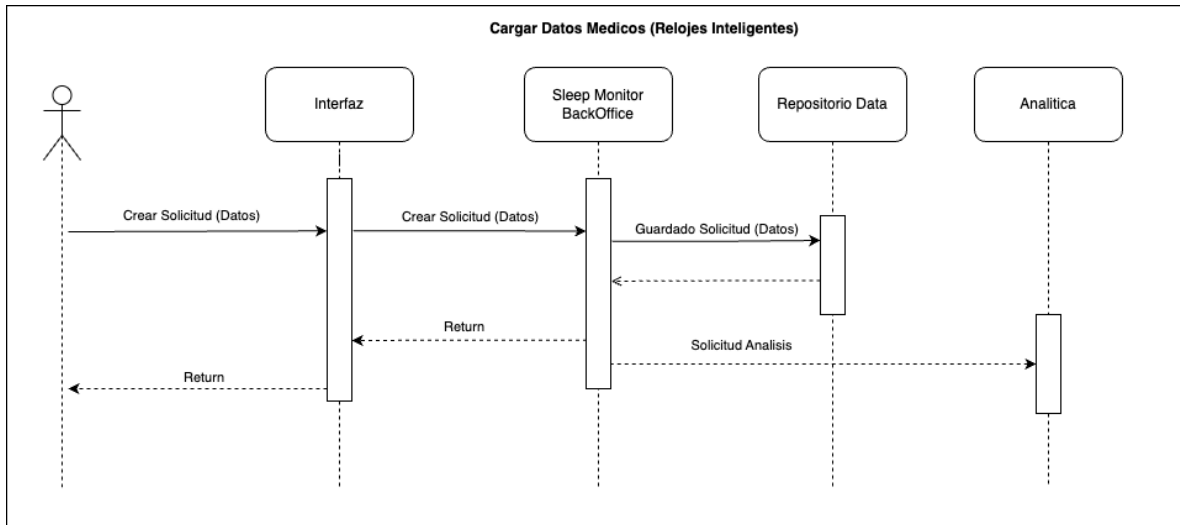


Ilustración 6 - Diagrama Secuencia Cargar Datos Médicos
Elaboración Propia

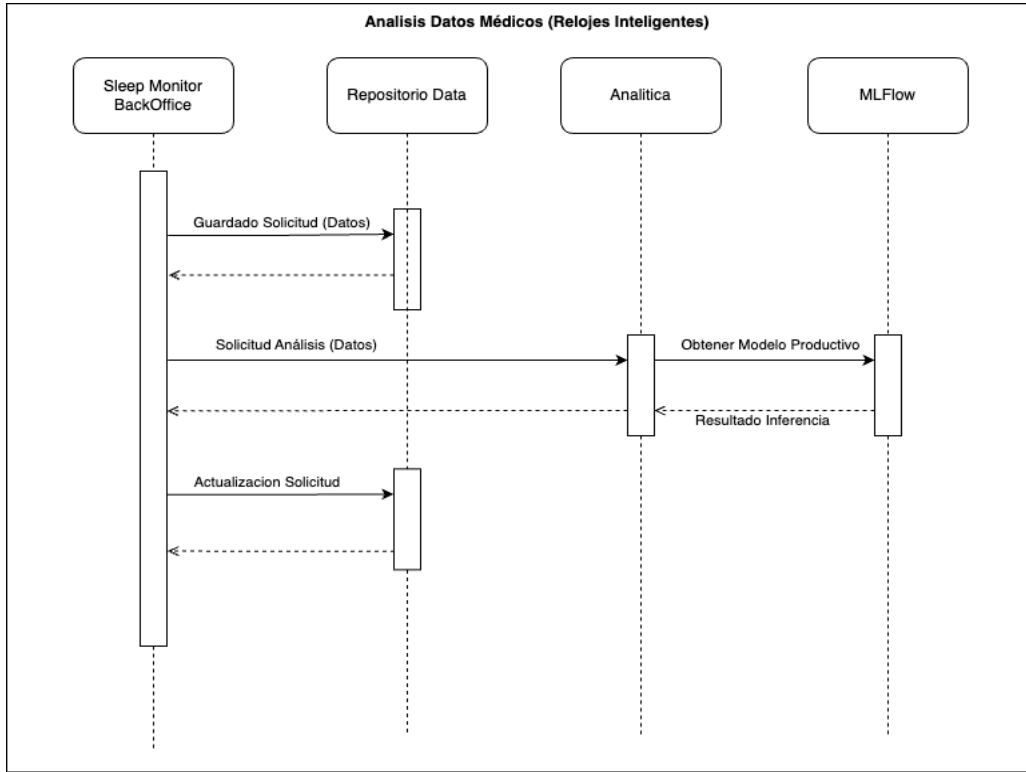


Ilustración 7 - Diagrama Secuencia Análisis de Datos Médicos
Elaboración Propia

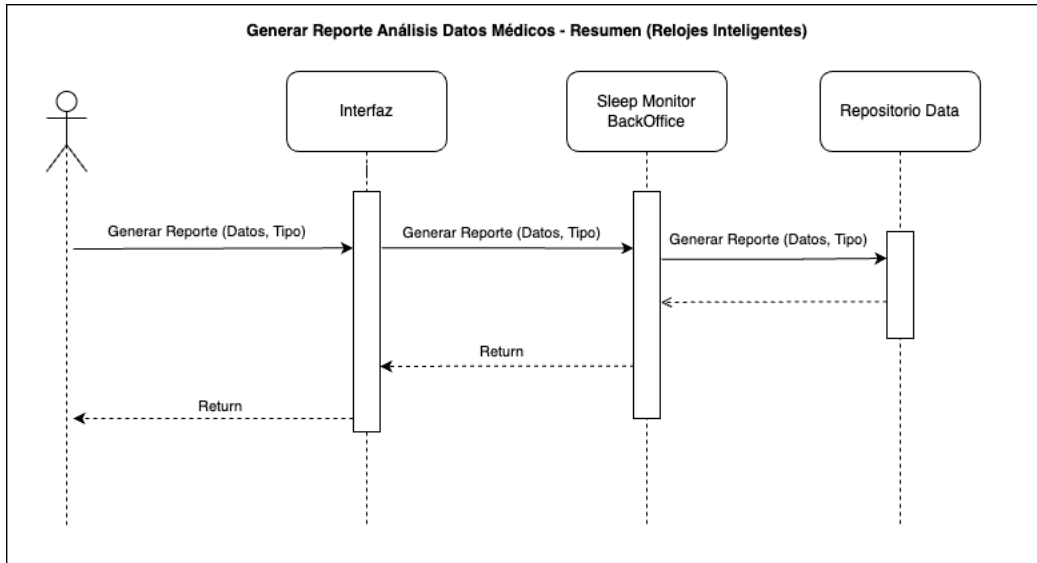
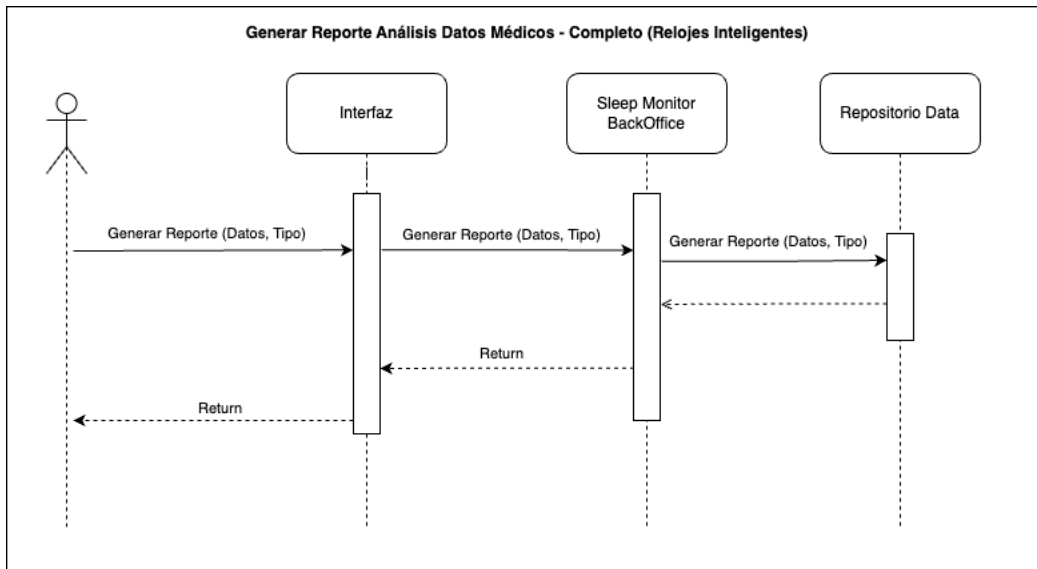
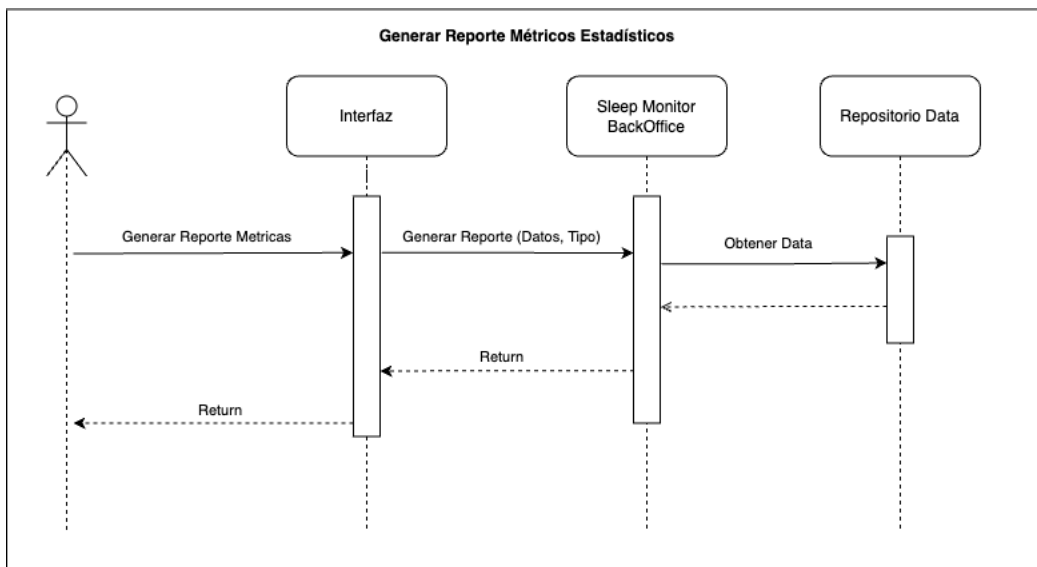


Ilustración 8 - Diagrama Secuencia Reporte Análisis Resumen
Elaboración Propia



*Ilustración 9 - Diagrama Secuencia Reporte Análisis Medico Completo
Elaboración Propia*



*Ilustración 10 - Diagrama Secuencia Reporte Métricas
Elaboración Propia*

4.5. Vista Despliegue

El funcionamiento del sistema implica la interacción de múltiples componentes distribuidos en diversos servidores y bases de datos. La "Aplicación Web" permitirá a los usuarios interactuar con el sistema a través de navegadores web, mientras que el "Servidor de Aplicación (BackOffice)" se encargará de gestionar usuarios, datos y presentar los resultados del análisis. Por otro lado, el "Servidor de Aplicación (Analítica)" será esencial para el entrenamiento del modelo y el análisis de datos, conectándose al "Servidor MLFlow" para el despliegue del modelo de inferencia. Para el almacenamiento de datos, se utilizarán una base de datos MySQL y una base de datos MongoDB (Datalake), ambas accesibles mediante protocolo TCP/IP. La coordinación efectiva de estos elementos en su puesta en funcionamiento es esencial para garantizar un sistema completo y robusto para el monitoreo y análisis del sueño de los usuarios.

El despliegue del sistema (ilustración 11) tiene enfoque on-premises proporcionando un mayor control sobre la seguridad y el acceso a los datos, garantizando un despliegue eficiente y escalable del sistema en un entorno de hardware dedicado. Esto se justifica principalmente debido a las restricciones y requerimientos específicos relacionados con la sensibilidad de los datos médicos. Dado que se trata de información altamente confidencial y sujeta a regulaciones de privacidad de datos, como HIPAA u otras normativas similares, mantener el control y la gestión interna de los datos es crucial para garantizar el cumplimiento normativo y la protección de la privacidad del usuario.

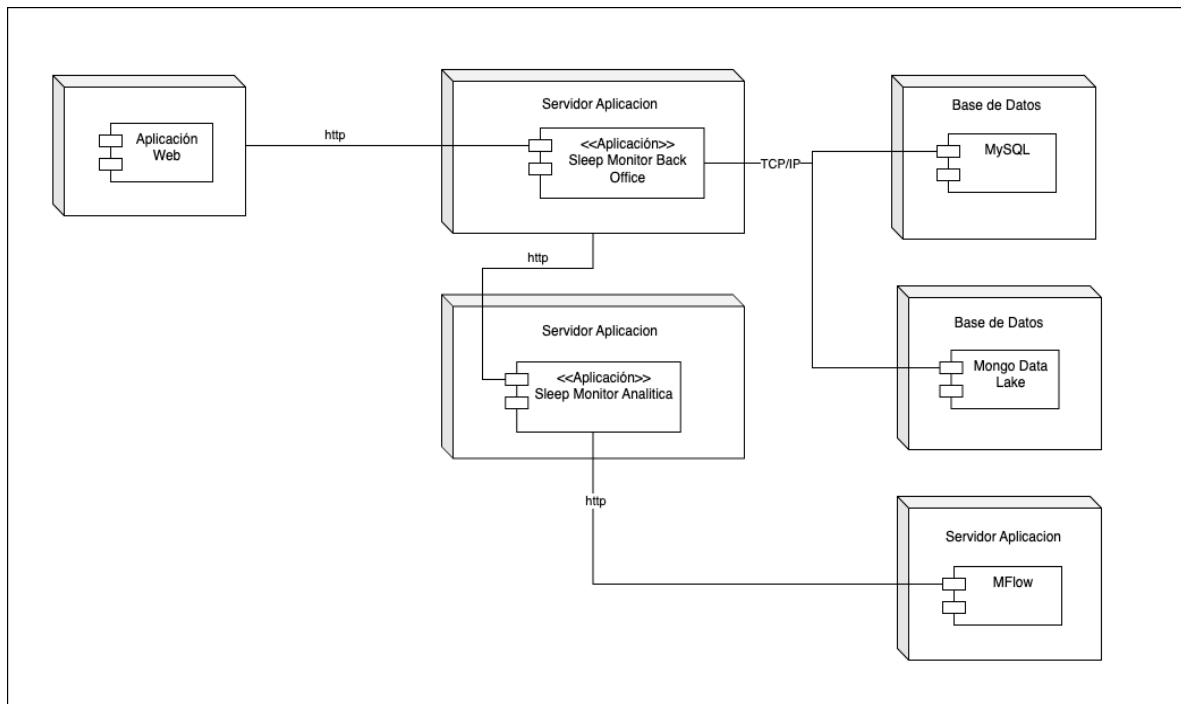


Ilustración 11 - Diagrama Despliegue
Elaboración Propia

5. Tamaño y Rendimiento

La arquitectura de software cumple con los requisitos de calidad, establecidos en el documento de Especificación de requisitos de software

- El sistema

6. Calidad

La arquitectura de software cumple con los requisitos de calidad, establecidos en el documento de Especificación de requisitos de software

- La interfaz de usuario del Sistema deberá estar diseñada contar una sencilla interacción y adecuada para una comunidad, sin necesidad de capacitación adicional en el sistema.
- El Sistema estará disponible las 24 horas del día, los 7 días de la semana, con un tiempo de inactividad no superior al 4%.
- El Sistema debe contar con una precisión de 90% en el modelo de inferencia