

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California).

La investigación fue financiada por el CONAHCYT (Consejo Nacional de Humanidades, Ciencias y Tecnologías).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos de Autor.

CICESE © 2023, Todos los Derechos Reservados, CICESE

Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California



Maestría en Ciencias en Ciencias de la Computación

Contratos inteligentes para la gestión de datos de sensado móvil y vestibular para aplicaciones en salud

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

José Ricardo Cedeño García

Ensenada, Baja California, México

2023

Tesis defendida por

José Ricardo Cedeño García

y aprobada por el siguiente Comité

Dr. Jesús Favela Vara
Director de tesis

Dr. Antonio García Macías

Dr. Roberto Conte Galván



Dr. Pedro Gilberto López Mariscal
Coordinador del Posgrado en Ciencias de la Computación

Dra. Ana Denise Re Araujo
Directora de Estudios de Posgrado

Resumen de la tesis que presenta José Ricardo Cedeño García como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Computación.

Contratos inteligentes para la gestión de datos de sensado móvil y vestible para aplicaciones en salud

Resumen aprobado por:

Dr. Jesús Favela Vara
Director de tesis

El aumento en la producción de datos derivado de la adopción de tecnologías móviles y de IoT está revolucionando la salud, pero también plantea importantes retos éticos y de privacidad. Los recientes avances en el aprendizaje automático han resaltado la importancia de recopilar y etiquetar datos correctamente, en especial para fines críticos, como el desarrollo de aplicaciones para cuidados médicos. La recopilación de datos médicos para tareas de aprendizaje automático presenta limitaciones en cuanto a la cantidad, variedad y calidad de las fuentes disponibles. Una forma de abordar este dilema es el uso de Blockchain para la recopilación y el uso de datos de pacientes. El anonimato de una red centralizada permite proteger la identidad del paciente. La estructura formada por nodos permite que la información esté siempre disponible y no dependa de un servidor principal. La inmutabilidad de los registros en la cadena garantiza la trazabilidad inequívoca del flujo de los datos del paciente. Por último, los mecanismos de consenso y recompensa de la red podrían motivar a nuevos usuarios a participar del sensado activo. Presentamos TRHEAD, una arquitectura de referencia basada en la Blockchain para recopilar datos sanitarios, firmar consentimientos, anotar datos y obtener crédito por los mismos, permitiendo a los usuarios rastrear el uso de sus datos, a los científicos rastrear su procedencia y proteger al mismo tiempo la privacidad de los pacientes. Exponemos dos implementaciones de nuestra arquitectura aplicadas a distintas campañas de sensado para comprobar su viabilidad, así como los resultados de su aplicación en estos escenarios y las conclusiones que desprendieron de su análisis. Dado que uno de los objetivos principales de TRHEAD es la recopilación de datos mediante sensado activo para el entrenamiento legal/consciente de modelos de aprendizaje automático, se realizó el entrenamiento de un modelo con los datos obtenidos de la campaña de sensado correspondiente a imágenes de rostros humanos, con el fin de detectar estados de ánimo. Finalmente se discute el papel de TRHEAD en el aseguramiento del trato justo y consciente de la información de los pacientes y el camino por recorrer en el perfeccionamiento de la arquitectura.

Palabras clave: Contratos Inteligentes, Blockchain, Privacidad, Aprendizaje de Máquina Ético, Recopilación Consciente de Datos, Consentimiento, Arquitectura de Referencia

Abstract of the thesis presented by José Ricardo Cedeño García as a partial requirement to obtain the Master of Science degree in Computer Science.

Smart contracts for mobile and wearable sensing data management for health applications

Abstract approved by:

PhD Jesús Favela Vara
Thesis Director

The increase in data production resulting from the adoption of mobile and IoT technologies is revolutionizing healthcare, but it also poses significant ethical and privacy challenges. Recent advances in machine learning have highlighted the importance of collecting and labeling data correctly, especially for critical purposes such as deploying healthcare software. Collecting medical data for machine learning tasks presents limitations in terms of the quantity, variety, and quality of available sources. One way to address this dilemma is the use of Blockchain for the collection and use of patient data. The anonymity of a centralized network allows the patient's identity to be protected. The structure formed by nodes allows information to be always available and not dependent on a main server. The immutability of the records in the chain guarantees the unequivocal traceability of the flow of patient data. Finally, the network's consensus and reward mechanisms could motivate new users to participate in active sensing. We present TRHEAD, a Blockchain-based reference architecture for collecting healthcare data, signing consents, annotating data and getting credit for it, allowing users to track the use of their data, scientists to track its provenance while protecting patients privacy. We present two implementations of our architecture applied to different sensing campaigns to test their feasibility, as well as the results of their application in these scenarios and the conclusions drawn from those results. Since one of the main objectives of TRHEAD is the collection of data through active sensing for the legal/conscious training of machine learning models, a model was trained with the data obtained from the sensing campaign corresponding to images of human faces, in order to detect moods. Finally, the role of TRHEAD in ensuring the fair and conscientious treatment of patient information and the road ahead in refining the architecture is discussed.

Keywords: Smart Contracts, Blockchain, Privacy, Ethical Machine Learning, Conscious Data Collection, Consent, Reference Architecture

Dedicatoria

A mi madre porque le debo todo lo que soy, a mi hermano porque siempre ha creído en mí y a todos los que me acompañaron en este camino.

Agradecimientos

Al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California, por la oportunidad de superarme realizando mis estudios de posgrado.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT) por brindarme el apoyo económico para realizar mis estudios de maestría. No. de becario: 1161287.

A mi Director de Tesis, Dr. Jesús Favela Vara, por su esfuerzo y dedicación constantes, su paciencia y los espacios en su ya apretada agenda que nunca me negó. Su ejemplo ha sido fundamental en mi formación como investigador. Gracias porque sé que por cada hora de sueño que sacrifiqué para mejorar mi trabajo, usted sacrificó dos.

A mi comité de Tesis, Dr. Antonio García Macías y al Dr. Roberto Conte Galván, por sus comentarios y observaciones que ayudaron a mejorar mi trabajo. A Carlos Eduardo Sánchez Torres por su excelente trabajo con la plataforma EvaNotebook y el consecuente entrenamiento y despliegue del modelo de aprendizaje de máquina.

Tabla de contenido

	Página
Resumen en español	ii
Resumen en inglés	iii
Dedicatoria	iv
Agradecimientos	v
Lista de figuras	viii
Lista de tablas	ix
Capítulo 1. Introducción	
1.1. Justificación	2
1.2. Pregunta de investigación	4
1.3. Objetivos	5
1.3.1. Objetivo general	5
1.3.2. Objetivos específicos	5
1.4. Metodología	6
Capítulo 2. Trabajo relacionado	
2.1. Mecanismos de consenso	7
2.1.1. PoW (Proof of Work)	8
2.1.2. PoS (Proof of Stake)	8
2.1.3. DPoS (Delegated Proof of Stake)	8
2.2. Aplicación de la Blockchain en la Salud	9
2.2.1. Contratos inteligentes	10
2.2.2. Tendencias	11
2.2.2.1. Salud Móvil	11
2.2.2.2. Inalámbrica	12
2.2.2.3. Internet of Things (IoT)	12
2.2.3. NFT	13
2.2.4. Tokenomics	13
2.2.5. NFT en los Servicios de Sanitarios	14
2.3. MLOps	14
Capítulo 3. La Arquitectura de Referencia TRHEAD	
3.1. Escenarios	15
3.1.1. Escenario 1.1 (Generador de Datos):	16
3.1.2. Escenario 1.2 (Tenedor de Datos):	16
3.1.3. Escenario 2 (Curador de Datos)	17
3.1.3.1. Tenedor de datos	17
3.1.3.2. Consumidor de datos	18
3.1.3.3. Curador de datos	18
3.1.3.4. Resultado	19
3.1.4. Escenario 3 (Consumidor de datos)	19
3.2. Funcionalidades necesarias en una plataforma de TRHEAD	20

3.3.	Componentes de TRHEAD	21
3.4.	DApp	22
3.5.	Sistema de almacenamiento en la nube	22
3.6.	Blockchain	22
3.7.	Modelo de Aprendizaje de Máquina	23
3.8.	Flujo de datos entre los componentes	24

Capítulo 4. Implementaciones de la arquitectura de referencia THREAD para el sentido móvil

4.1.	Componentes	26
4.1.1.	Interfaz gráfica	26
4.1.1.1.	Subir datos de salud	27
4.1.1.2.	Sensado de datos de acelerometría	28
4.1.1.3.	Listado de activos	29
4.1.2.	Servidor Web	30
4.1.3.	Contrato Inteligente y la Blockchain	30
4.1.4.	Aplicación externa	32
4.2.	Pruebas	32
4.2.1.	Procedimiento	33
4.2.2.	Encuestas de experiencia de usuario	33
4.2.3.	Resultados de la encuesta	34
4.2.3.1.	Cuestionario de intención de uso	34
4.2.3.2.	Cuestionario de aceptación de la tecnología (TAM)	35
4.2.4.	Cuestionario de conformidad con la tecnología	37
4.3.	Conclusiones parciales del experimento	37
4.4.	La implementación FRED	39
4.4.1.	Interfaz móvil	39
4.4.1.1.	Captura de imágenes	40
4.4.1.2.	Vista de validación de imágenes	40
4.4.1.3.	Listado de imágenes	41
4.4.2.	Contrato inteligente	41
4.4.3.	Entrenamiento de modelos predictivos	41
4.4.3.1.	Afinamiento	42
4.4.3.2.	Despliegue en EvaNotebook	47

Capítulo 5. Discusión

5.1.	Desafíos técnicos	48
5.2.	Desafíos Éticos	50
5.2.1.	Uso justo de los datos	50
5.2.2.	Protección de la privacidad	51
5.2.3.	Sesgo	51
5.2.4.	Cultura de la justeza	52

Capítulo 6. Conclusiones

6.1.	Contribuciones	54
6.2.	Trabajo a futuro	55

Literatura citada	56
-----------------------------	----

Lista de figuras

Figura	Página
1. Estructura de la arquitectura de referencia TRHEAD	21
2. Flujo de datos entre los componentes del Escenario 1.1	25
3. Interfaz de gestión de archivos	27
4. Interfaz de captura de datos de acelerometría	28
5. Listado de activos en la plataforma	29
6. Código del contrato inteligente “Consent Letter” escrito en el lenguaje Solidity	31
7. Flujo de las direcciones de los datos para su consumo final	32
8. Resultados de los cuestionarios de Intención de Uso	35
9. Resultados de los cuestionarios de TAM, sección de utilidad	36
10. Resultados de los cuestionarios de TAM, sección de usabilidad	36
11. Captura de imágenes	40
12. Validación de imágenes	40
13. Lista de imágenes	40
14. Contrato inteligente para manejar metadatos de imágenes clasificadas	42
15. Distribución de las clases en el conjunto de imágenes	43
16. F1-Score del modelo entrenado	44
17. AUC-PR de la curva de “Precisión-Recuperación” que indica la capacidad de clasificación del modelo.	45
18. Matriz de confusión normalizada	45
19. Ejemplo de detección de gestos faciales por el modelo	46

Lista de tablas

Tabla	Página
1. Respuestas a la encuesta de conformidad con la tecnología	38
2. Derechos y deberes de los usuarios según su rol	53

Capítulo 1. Introducción

La proliferación de los dispositivos móviles, vestibles y de Internet de las Cosas (IoT) ha permitido su introducción en nuestra vida diaria a un nivel sin precedentes (Ray et al., 2019). Los sistemas de salud han sabido aprovechar estos avances para ofrecer sus servicios de forma remota y con una interrupción mínima de las actividades diarias del paciente. Surge así la salud electrónica o e-Health (Eysenbach, 2001) con el fin de mejorar el acceso, la eficiencia y la calidad de los cuidados personalizados de salud; a través de la fusión de dispositivos de sensado vestibles y sistemas digitales en Internet. Los dispositivos móviles, por ejemplo, son en sí mismos un arreglo de sensores que recolectan datos de ubicación, audio, imagen, texto, entre otros; de los cuales se puede inferir información adicional mediante técnicas de reconocimiento de actividades (Kwapisz et al., 2011).

Una de las principales aplicaciones del cómputo ubicuo se encuentra en el área de la salud y el bienestar, aprovechando la capacidad de estos dispositivos de expandir el acceso a los servicios, reducir costos y su gran adopción por parte del consumidor; para monitoreo de salud a distancia, sistemas inteligentes de manejo de emergencias y acceso a datos sanitarios (Bardram, 2008). Esto se conoce como “pervasive health” o “salud penetrante” y ha atraído la atención de los proveedores de servicios sanitarios, provocando la aparición de la m-Health (Kotz et al., 2016). La digitalización de los servicios de salud ha llevado a un incremento en el volumen y tipo de datos de salud que se recopilan, así como de los análisis que se pueden llevar a cabo con estos (Mann et al., 2021). Se espera que este volumen de información se haga aún mayor en la medida en que los pacientes están más dispuestos a participar en la toma de decisiones sobre su salud; pero desafortunadamente los dispositivos que recolectan esta cantidad masiva de datos implementan medidas de seguridad ligeras al ser de bajos recursos (Zhang et al., 2016). Estos sensores tienden a formar parte de una arquitectura cliente-servidor donde todos los datos recolectados se centralizan en una base de datos y se envían o consultan a través del protocolo HTTP.

Por este motivo los servicios digitalizados de salud adolecen de muchos de los problemas que los amplios sistemas de servidores centralizados sufren: el intercambio de datos y la gestión del consentimiento, el control de acceso, la autenticación y la confianza del usuario (Kotz et al., 2016). Una falla en la seguridad de estos sistemas podría dar paso a un ataque cibernético que comprometa la información personal del usuario haciéndolo víctima del robo de identidad o de un incidente que ponga en riesgo su salud por pérdida de información crítica para un correcto dictamen médico. Los riesgos van más allá de factores técnicos, las acciones de empresas deshonestas provocan directamente la fuga, uso incorrecto o

comercialización de los datos del paciente con un fin lucrativo ¹.

A pesar de que la solución más lógica podría parecer la prohibición del uso de datos de los pacientes o usuarios en general, debemos entender la importancia crítica de los mismos en el mejoramiento de los servicios de nuestra sociedad. El enfoque correcto es el de generar “datos justos”, que cumplen con ser **localizables, accesibles, interoperables y reusables** (FAIR, por sus siglas en Inglés) (Dunning et al., 1970) y el uso justo de los mismos; cumpliendo con los principios de **beneficio colectivo, control de autoridad, responsabilidad y ética** (CARE, por sus siglas en Inglés)(Leonelli et al., 2021). En este trabajo proponemos el uso de la tecnología Blockchain por su capacidad de funcionar como registro de transacciones y almacenamiento de información descentralizado, distribuido y democrático. Su estructura basada en nodos permite que las funciones de la plataforma como el registro y almacenamiento de datos se compartan entre los participantes de la misma, de manera que se distribuya la carga de trabajo, prescindiendo de un servidor principal. En cada nodo existe una copia del estado de la red, por lo que si uno de ellos falla, la información está disponible en el resto. Esta estructura participativa permite que no sea una institución o persona jurídica quien tome las decisiones, sino que se llega a un consenso para la satisfacción de la mayoría.

La finalidad de este trabajo es la presentación de una arquitectura de referencia (Angelov et al., 2012) para la recopilación y uso justos de los datos de sensado móvil y vestible. Bajo el nombre Traceable Health Data (TRHEAD) planteamos una arquitectura para sistemas basados en la Blockchain capaz de recolectar datos de sensado a la vez que se respeta la privacidad del usuario, provee control y autoría sobre los mismos registrando cada operación en los bloques de la cadena y garantizando datos conscientemente recopilados para el entrenamiento de modelos de aprendizaje automático.

1.1. Justificación

Los gobiernos necesitan datos de gastos médicos de los contribuyentes para la planeación del presupuesto de salud y a menudo estos datos se encuentran esparcidos entre distintos proveedores de servicios sanitarios que difieren en procedimientos de captura y metas; provocando una brecha entre las necesidades

¹El 4 de marzo, la FTC (Comisión Federal de Comercio de los Estados Unidos) emitió una orden de conciliación contra WW International (la empresa anteriormente conocida como Weight Watchers) y su filial, Kurbo. La presidenta de la FTC, Lina Khan, declaró que “Weight Watchers y Kurbo comercializaron servicios de control de peso para su uso por niños de tan sólo ocho años, y luego recopilaron ilegalmente información personal y sensible sobre su salud”.

reales y las percibidas por los gobiernos (Zghaibeh et al., 2020). Los datos epidemiológicos se podrían usar para determinar patrones de conducta en los pacientes que puedan ser la causa de las enfermedades que los aquejan (Favela et al., 2016). Desórdenes de salud que parecen aislados podrían estar relacionados con una afección más grave cuyos síntomas aún no son visibles y convertirse en una señal de alarma para quienes los padecen. Para establecer esta relación es necesario recabar datos de estos pacientes para investigar el clúster de síntomas que sufren y plantear un tratamiento acorde (Fiorentino et al., 2011). En general los estudios clínicos que sirven para mejorar la atención médica necesitan de biomarcadores específicos para poder asegurar que la misma será efectiva para todos, sin importar etnia, edad o sexo.

Habiendo expuesto la necesidad del consumo de datos médicos para el mejoramiento de los sistemas de servicios sanitarios y los peligros que un mal manejo de los mismos puede acarrear, se hace evidente la necesidad de contar con una plataforma capaz de controlar el flujo de datos médicos proveniente del sensado móvil y vestible bajo ciertos principios.

La **trazabilidad** de los datos, por ejemplo, permite a los consumidores de datos conocer el recorrido que estos han cursado antes de llegar a un dataset, base de datos o repositorio; permitiendo corroborar que no han sido alterados o su integridad ha sido comprometida.

La **proveniencia** de los datos permite conocer que estos se recopilieron de forma legal, se etiquetaron correctamente y no han sido generados de forma sintética (Shumailov et al., 2023). Sin embargo, estos principios a menudo implican procesos de investigación que comprometen la identidad del paciente que los generó o permiten determinar información del mismo que no está permitido revelar. Como ejemplo, la recopilación de datos de una banda inteligente puede ofrecer información de geolocalización que al ser sometida a escrutinio manual o de un programa informático puede revelar la residencia y lugares que frecuenta el usuario del sensor (Kalnis et al., 2007).

Entonces, otro principio debería ser el **respeto a la privacidad** del paciente, que no se debe confundir con la protección de sus datos, puesto que a pesar de hacerlos públicos para su consumo, se podría remover cualquier conexión con la identidad de quien los generó. Otra opción es la creación de una capa de anonimato entre el paciente y sus datos, una identidad electrónica aislada de la información de la persona física hacia la cual enfocar la trazabilidad y proveniencia de los mismos.

El **consentimiento** es imprescindible en este escenario y comprende el derecho del paciente a con-

trolar y conocer la forma en que sus datos están siendo usados, así como el tiempo de vida de los mismos. Cuando se recolectan datos para un estudio clínico se debe respetar la decisión de negar o revocar el consentimiento de su uso y no debería este ser la condición para la prestación de un servicio si no es estrictamente necesario, según la Regulación General para la Protección de los Datos (GDPR).²

La **disponibilidad** responde a la necesidad de contar con los datos en todo momento e implica que el servicio que los provee no se detenga por factores técnicos o se reduzca por no poder responder a la demanda. Viene ligada a la **ausencia de sesgo** que significa contar, no solo con cantidad, sino con variedad de datos que incluyan representaciones de todas las clases a las que va dirigido un producto. Esto permite que productos como los algoritmos de aprendizaje de máquina sean entrenados sin sesgo y respondan de manera acertada, en vez de inclinar sus respuestas hacia una clase dominante del conjunto de datos de entrenamiento. Para cumplir con estos principios la plataforma debe estar soportada en una tecnología:

- **Transparente:** donde se puedan asegurar la **trazabilidad** y la **proveniencia** de los datos.
- **Participativa:** donde un usuario pueda defender el derecho al respeto de su **privacidad** y al **consentimiento**.
- **Decentralizada:** con el fin de que la información esté **disponible** desde varios proveedores y la plataforma alcance a más usuarios, aumentando a la vez la **variedad de los datos** que se recopilan.

1.2. Pregunta de investigación

Considerando los requisitos antes expuestos para el manejo de información de sensado móvil y vestible destinada al entrenamiento de modelos de aprendizaje de máquina, planteamos la siguiente pregunta de investigación:

¿En qué medida el uso de contratos inteligentes permite atender algunas de las limitaciones de trazabilidad y no refutabilidad asociadas al manejo de datos de sensado móvil para su consumo en el entrenamiento consciente/legal de modelos de aprendizaje de máquina?

²La GDPR afirma ser una de las leyes de seguridad y privacidad más abarcadoras del mundo. Impone obligaciones a empresas que se dirijan a personas de la Unión Europea o recojan datos relacionados con ellas. El incumplimiento de estas obligaciones acarrea multas que pueden alcanzar las decenas de millones de euros

1.3. Objetivos

1.3.1. Objetivo general

Proponer una arquitectura de referencia para sistemas de recolección y uso de datos de sensado móvil y vestible, basado en contratos inteligentes que se ejecutan sobre la red Blockchain; aprovechando las capacidades de la “cadena de bloques” para ofrecer la seguridad, disponibilidad, procedencia e integridad de los datos a la vez que se asegura la trazabilidad y no refutabilidad de las operaciones sobre estos sin comprometer la privacidad del usuario que los publicó.

1.3.2. Objetivos específicos

Para cumplir con el objetivo general planteado se proponen los siguientes objetivos específicos:

- Proponer y documentar una arquitectura de referencia (TRHEAD) que sirva como directriz en la implementación de sistemas que busquen cumplir con lo expuesto en 1.3.1.
- Implementar un sistema que cumpla con los requisitos de TRHEAD para medir la factibilidad de su despliegue.
- Desarrollar contratos inteligentes que se ejecuten sobre la red Blockchain para automatizar el control de acceso, edición y consumo de los datos.
- Desarrollar una interfaz fácil de usar que permita a los usuarios no técnicos interactuar con la red de bloques.
- Desarrollar una Interfaz de Programación de Aplicaciones (API) para que los datos manejados en el sistema puedan ser consumidos desde una herramienta externa para el entrenamiento de modelos de aprendizaje de máquina conscientemente/legalmente entrenados.
- Realizar una prueba con un grupo de sujetos reducido para medir la viabilidad del uso del sistema en un ambiente real.
- Discutir las ventajas y desventajas de la arquitectura propuesta con respecto a otras no basadas en la Blockchain y sus implicaciones éticas.

1.4. Metodología

Con el fin de cumplir los objetivos planteados se siguió una metodología de trabajo que incluye el desarrollo de la arquitectura de referencia TRHEAD y la implementación de un sistema a partir de las pautas que esta comprende. Previa al desarrollo del sistema, se realiza una investigación con el fin de conocer las tecnologías más recomendables para la construcción de una interfaz fácil de usar y compatible con una amplia gama de dispositivos, capaz de desplegar nuevos contratos inteligentes e interactuar con ellos. Posteriormente se realizaron experimentos para medir la viabilidad de la implementación de esta arquitectura con un grupo reducido de usuarios con un trasfondo asiduo a la ciencia de datos. Una vez concluidos los experimentos se aplicaron cuestionarios para medir la usabilidad, aceptación e intención de uso de los usuarios para con el sistema.

A continuación, en el Capítulo 2, se exponen los antecedentes que sirvieron de punto de partida para la propuesta. En el Capítulo 3 se expone la arquitectura de referencia arriba mencionada y los escenarios a los que está enfocada. Llegado el Capítulo 4 se muestran ejemplos de implementación de la arquitectura y las pruebas realizadas para medir su viabilidad en los escenarios planteados. El Capítulo 5 está dedicado a la discusión de los resultados obtenidos desde un punto de vista técnico y ético. Finalmente en el Capítulo 6 se arribarán a conclusiones del trabajo investigativo y de implementación realizado.

Capítulo 2. Trabajo relacionado

La tecnología de registro contable distribuida (DLT) se ha establecido como un término que abarca los sistemas distribuidos que operan en un ambiente sin un operador central o autoridad, a pesar de los participantes que pueden ser poco confiables o de intenciones maliciosas (ambiente adversario). Blockchain o “red de bloques” es una tecnología a menudo considerada como una especificación de las DLT con una estructura de datos particular que consiste de una cadena de bloques de datos enlazados por funciones “hash” (Rauchs et al., 2018). Las funciones “hash” son técnicas de criptografía que permiten crear mediante una ecuación y una clave de entrada, un dato que será una posición en un arreglo, se conoce como “hash” al valor que se genera la función al procesar una entrada que se desea encriptar.

Las transacciones son replicadas (distribuidas) en copias idénticas entre cada uno de los usuarios de la red. La nueva información solo es agregada al registro principal por consenso, cuando todos los usuarios están de acuerdo en que la información es precisa. Cualquier intento de alterar la información por un solo usuario será transparente al resto, asegurando, en teoría, la inmutabilidad de las transacciones (Pearson et al., 2019). Blockchain es una DLT donde cada nueva información que se desea registrar se verifica y se vincula al bloque anterior, formando una cadena larga. Como cada transacción es registrada y verificada públicamente, proporciona una gran confiabilidad. Al ingresar un registro, nadie puede modificar la información escrita en la cadena de bloques. Sirve para demostrar que los datos son reales y no han sido manipulados. Los datos se mantienen en la red en lugar de una base de datos central, mejorando la estabilidad y resistencia a ser atacada.

2.1. Mecanismos de consenso

En los sistemas distribuidos no existe un protocolo de consenso perfecto. El protocolo de consenso debe establecer un equilibrio entre la coherencia, la disponibilidad y la Tolerancia a Fallos de Partición (CAP). También debe abordar el Problema de los Generales Bizantinos, ya que habrá nodos maliciosos que socaven deliberadamente el proceso de consenso. En esta sección, hacemos una descripción detallada de algunos protocolos de consenso de Blockchain (Zhang et al., 2016).

2.1.1. PoW (Proof of Work)

Para la Prueba de Trabajo o de Participación se selecciona un nodo para crear un nuevo nuevo bloque en cada ronda de consenso mediante una competición. En la competición, los nodos participantes deben resolver un reto criptográfico. El nodo que primero resuelva el reto puede tener derecho a crear un nuevo bloque. Es, computacionalmente hablando, muy costoso resolver un reto de PoW. Los nodos tienen que ajustar un valor conocido como “nonce” (number used once) ¹ para obtener la respuesta correcta, lo que requiere mucha potencia de cálculo. A medida que aumentan los bloques válidos en la cadena, la carga de trabajo también se acumula, por lo que derrocar una cadena larga requiere cada vez mayor cantidad de potencia de cómputo. PoW pertenece a los protocolos de consenso de finitud probabilística ya que garantiza la consistencia final.

2.1.2. PoS (Proof of Stake)

En PoS, la selección en cada ronda del nodo que crea un nuevo bloque depende de la participación mantenida en lugar de la potencia computacional. Aunque los nodos siguen necesitando resolver un rompecabezas criptográfico, la diferencia con PoW es que los nodos no necesitan ajustar el valor “nonce” muchas veces, en cambio, la clave para resolver este rompecabezas es la cantidad de “stake” (participación o cantidad de tokens bloqueados por ronda). Por lo tanto, PoS es un protocolo de consenso que ahorra energía, aprovechando el incentivo interno de la moneda en lugar de consumir mucha potencia computacional para alcanzar un consenso. PoS también es un protocolo de consenso de finitud probabilística.

2.1.3. DPoS (Delegated Proof of Stake)

El principio de DPoS es dejar que los nodos que tienen participación voten para elegir a los verificadores de bloques (es decir, creadores de bloques). Esta forma de votación hace que quienes poseen “stake” den el derecho de crear bloques a los delegados que apoyan, en lugar de crearlos ellos mismos, reduciendo de esta manera el consumo de poder computacional a cero. DPoS es como un sistema parlamentario, si

¹Número que se usa solo una vez en un cálculo o proceso criptográfico específico.

los delegados no pueden generar bloques en sus turnos, serán destituidos y los interesados seleccionarán nuevos nodos para sustituirlos. DPoS aprovecha al máximo los votos de los accionistas para alcanzar un consenso de forma justa y democrática. En comparación con PoW y PoS, DPoS es un protocolo de consenso de bajo coste y alta eficiencia.

2.2. Aplicación de la Blockchain en la Salud

En el ámbito de la salud, el uso de la Blockchain permite mantener un registro general e inmutable de los procedimientos médicos y los cuidados del paciente, disminuyendo la cantidad de prácticas médicas y monitoreo necesarios cuando el paciente es trasladado a otra institución. Esto permite ahorrar tiempo y recursos tanto a practicantes como a pacientes; estos últimos podrán ver a dónde va su información con total transparencia. Las firmas farmacéuticas podrán deshacerse de los problemas de manipulación y falsificación de los medicamentos al realizar un rastreo de cada uno de sus activos (Haleem et al., 2021). Siendo una tecnología relativamente nueva, la Blockchain no ha permeado de una manera notable fuera del ámbito de las TIC. A menudo se comete el error de etiquetarla sólo como la base para el auge que están teniendo las cripto-divisas. Sin embargo, sus prestaciones responden a muchas de las necesidades de privacidad, disponibilidad, autenticidad, seguridad y trazabilidad del proceso de manejo de la información en instituciones de salud. El objetivo de las secciones a continuación es el de dibujar el panorama en que se encuentra la aplicación de las DLT, específicamente la Blockchain, en los sistemas de salud.

La red por sí sola no es capaz de manejar los procesos internos de las instituciones de salud, para ello existen aplicaciones que se construyen sobre esta red conocidos como contratos inteligentes: Como cualquier contrato, establece los términos de un acuerdo. Pero a diferencia de uno tradicional, los términos se ejecutan como código en una cadena de bloques. Los contratos inteligentes permiten a los desarrolladores crear aplicaciones que se benefician de la seguridad, la fiabilidad y la accesibilidad de la cadena de bloques, al tiempo que ofrecen sofisticadas funciones entre pares, desde préstamos y seguros hasta logística y juegos, según el sitio oficial de Coinbase.

2.2.1. Contratos inteligentes

Un contrato inteligente es un código que se ejecuta en la cadena de bloques para facilitar, ejecutar y hacer cumplir los términos de un acuerdo entre partes no confiables. Puede considerarse como un sistema que libera activos digitales a todas o algunas de las partes implicadas una vez que se han cumplido las reglas predefinidas. En comparación con los contratos tradicionales, los contratos inteligentes no dependen de un tercero de confianza para funcionar, lo que da lugar a bajos costes de transacción. Existen diferentes plataformas de “cadena de bloques” que pueden utilizarse para desplegar contratos inteligentes, siendo Ethereum una de las que más ha proliferado. Esto se debe a que el lenguaje Solidity soporta la característica de completitud de Turing que permite crear contratos personalizados. Los contratos inteligentes pueden aplicarse en manejo de activos, comercio electrónico y gestión de derechos musicales, según (Alharby & van Moorsel, 2017).

Estos contratos permiten programar diferentes funcionalidades y construir diversos servicios. Actúan como agentes de software inteligentes para administrar automáticamente determinadas transacciones cuando se cumplen las condiciones preasignadas en una red de cadenas de bloques, lo que los hace adecuados para los procesos internos automatizables, ya que eliminan la necesidad de la interferencia de terceros (Omar et al., 2020).

En el 2018, (Hölbl et al., 2018) describe la construcción de redes con algunas restricciones para los participantes (redes de consorcio y privadas), para dar a la red un mejor control sobre el acceso y la creación de registros. Las dos plataformas más utilizadas en ese momento eran Ethereum e Hyperledger Fabric, ambas compatibles con los contratos inteligentes. Los contratos inteligentes añaden muchas funcionalidades a una cadena de bloques, pero los investigadores no parecen utilizarlos con tanta frecuencia como cabría esperar.

Las tendencias actuales de la investigación sobre la cadena de bloques en la salud indican que se utiliza principalmente para compartir datos, registros sanitarios y control de acceso, pero rara vez para otros escenarios, como la gestión de la cadena de suministro o la gestión de la prescripción de medicamentos. Por lo tanto, el potencial de la cadena de bloques sigue sin explotarse. La mayoría de las investigaciones presentan un marco, una arquitectura o un modelo novedosos que utilizan la tecnología blockchain en los sistemas de salud. Además, a menudo no se ofrecen detalles técnicos sobre los elementos de la cadena de bloques utilizados, como la plataforma, el algoritmo de consenso, el tipo (pública, privada, consorcio)

o el uso de contratos inteligentes. La mayoría de las investigaciones también podrían proporcionar un prototipo de implementación o, al menos, discutir algunos detalles de implementación de sus propuestas.

Para el año 2019, según lo descrito en (McGhin et al., 2019), existe una visión más clara del lugar de la red de bloques en la “industria de la salud”. Viene a jugar un papel importante en la recopilación de datos médicos de los pacientes ya sea mediante digitalización o importación de sus expedientes. En este proceso funciona como plataforma de gestión de privilegios y almacenamiento, controlando el acceso a estos datos y facilitando las auditorías. Resuelve el problema de la interoperabilidad al reunir toda la información en una misma plataforma que utiliza el mismo método de encriptación, lo que facilita la migración de datos sin el peligro de perder el acceso a la misma o verse en la necesidad de adaptar su formato.

2.2.2. Tendencias

La movilidad es un requisito cada vez más importante en el sector de la salud, los pacientes son más propensos a usar dispositivos de sensado vestibles y exigen que sus registros tengan el mismo nivel de portabilidad. Los dispositivos inteligentes, los sensores y otros dispositivos dependientes de la Internet, ganan terreno y surge la necesidad de transportar los datos que generan. La necesidad de compartir y acceder a estos datos en tiempo real, desde cualquier lugar y en cualquier dispositivo, agrava el reto de garantizar que sean seguros y estén protegidos. El concepto de movilidad se divide en tres secciones principales: Salud Móvil, Inalámbrica e IoT.

2.2.2.1. Salud Móvil

La salud móvil (mHealth) es un campo en expansión en las aplicaciones sanitarias que incluye dispositivos como sensores miniaturizados, redes inalámbricas de baja potencia y teléfonos inteligentes. La mHealth adolece de muchos de los problemas que los amplios sistemas de servidores centralizados sufren: el intercambio de datos y la gestión del consentimiento, el control de acceso, la autenticación y la confianza del usuario (Kotz et al., 2016). Lograr la privacidad y la seguridad en una red de sensores inalámbricos (WSN) o de IoT requiere amplios recursos, mientras la información sanitaria mal gestionada y compro-

metida puede perjudicar al paciente y las perspectivas de futuro de las aplicaciones sanitarias móviles (Sahi et al., 2017). Estas amenazas surgen debido a que las organizaciones sanitarias que utilizan estas tecnologías pueden carecer de los conocimientos necesarios para proteger adecuadamente los datos de los pacientes (Kotz et al., 2016).

2.2.2.2. Inalámbrica

El despliegue de una red de área corporal inalámbrica (WBAN) puede incluir sensores corporales portátiles, y las amenazas de seguridad relacionadas incluyen la integridad de los datos, la disponibilidad de la red, la autenticación de los datos, la gestión segura, fiabilidad, localización segura, responsabilidad y flexibilidad (Al-Janabi et al., 2017). Los dispositivos de una WBAN también son de bajos recursos (Zhang et al., 2016), por lo que se necesitarán soluciones de seguridad ligeras.

2.2.2.3. Internet of Things (IoT)

La prevalencia de la tecnología IoT en el área de la salud está creciendo a medida que los pacientes están más dispuestos a participar en la toma de decisiones sobre su salud (Zhang et al., 2016). Para ello se utilizan dispositivos y sensores inteligentes que registran y envían datos vitales a su médico para observar y evaluar a distancia las condiciones crónicas. Ejemplos de estas tecnologías son las lentes de contacto, las bandas de ejercitación, los microchips bajo la piel y sensores inalámbricos (Torre et al., 2016). Estos sistemas inalámbricos a veces no tienen tan en cuenta la seguridad como otros sistemas más sensibles, lo que los expone a ataques de violación de seguridad.

En este contexto existen varias aplicaciones para las propiedades de la Blockchain, como el almacenamiento descentralizado, que consiste en aprovechar el carácter descentralizado de esta red para distribuir la información entre varios servidores. Esto permite un acceso más rápido a la información, interoperabilidad entre distintas instituciones de salud y el consecuente aumento de la cantidad y calidad de los datos para estudios médicos.

2.2.3. NFT

Un token (ficha) no fungible (NFT) es un “activo criptográfico en una blockchain que contiene información identificativa única y códigos que los separan unos de otros” según (Taherdoost, 2022). Aunque CryptoKitties, un juego basado en la red Ethereum atrajo inicialmente la atención sobre los bienes NFT a finales de 2017, su rápida expansión está íntimamente ligada al nacimiento del COVID-19. Después de que la Organización Mundial de la Salud identificó una epidemia global, la liquidez del mercado Bitcoin saltó dramáticamente. De manera significativa, las medidas de bloqueo de la pandemia aumentaron la interacción en Internet e impulsaron aún más la expansión de los NFT. Sin embargo, se ha investigado poco sobre estos en comparación con la extensa literatura sobre Bitcoin y otras criptomonedas.

2.2.4. Tokenomics

La tokenómica (tokenomics) es un sistema económico basado en la tokenización que comprende las características de los tokens, la política monetaria y los mecanismos de incentivación de los usuarios. En la tokenómica, los participantes reciben fichas (tokens) como compensación por su labor y las utilizan para comprar bienes y servicios. El término “tokenómica” se refiere al comportamiento económico del método de asignación de fichas en comparación con el mecanismo de asignación de recursos de la economía de mercado. Existen tres categorías principales de fichas en función de su uso previsto. El primer tipo incluye los tokens de pago, también conocidos como monedas, dígame Bitcoin y Ethereum. A diferencia de los “Fungible Tokens” (FT), que son distintos y divisibles, estos tokens no se adhieren al principio NFT.

Los tokens de utilidad son el segundo tipo de token, y están destinados a proporcionar a los usuarios acceso digital a una determinada aplicación o servicio. El tercer tipo, los tokens de activos, ha experimentado una enorme expansión en su uso gracias a la llegada de los NFT. Los estudios de Tokenomics se han centrado hasta ahora en el desarrollo, la distribución y la gestión de los ecosistemas de tokens según (Taherdoost, 2022). Los autores (Howell et al., 2020) sostienen que, si se lleva a cabo correctamente, la emisión de tokens tiene el potencial de superar a los instrumentos financieros convencionales en términos de seguridad, liquidez y transparencia.

2.2.5. NFT en los Servicios de Sanitarios

Con el fin de abordar los problemas y limitaciones existentes en los sistemas sanitarios en cuanto al almacenamiento, intercambio y acceso a los datos, (Vijayalakshmi et al., 2022) proponen una solución basada en NFTs. Implica la creación de una aplicación móvil para el almacenamiento de los expedientes médicos de los pacientes usando la tecnología Blockchain. El uso de esta herramienta provee seguridad de los datos almacenados y el seguimiento de las consultas que se hacen a los mismos. Se plantea que los exámenes médicos como escáneres, radiografías y resultados de laboratorio sean almacenados por la propia institución de la salud en el expediente digital del paciente. Este sería un paso de avance en vías de dar al usuario el control de su información sanitaria, especialmente en cuanto a quién tiene acceso a la misma. La importancia de este control radica en la necesidad de estos datos y su correcta manipulación en la realización de tareas de MLOps.

2.3. MLOps

Es un conjunto de principios y prácticas adoptados de DevOps y aplicados al aprendizaje automático. Se conocen tres tipos diferentes de configuraciones de MLOps con diversos grados de automatización: implementación manual, entrega continua de modelos e integración continua/entrega continua de canalizaciones (Sridhar et al., 2021). Surge con el fin de desplegar modelos de aprendizaje automático rápidamente y mantenerlos una vez desplegados. Tiene en cuenta todos los factores necesarios para la confección de un modelo funcional que requiere importantes recursos para planificar, desarrollar y probar.

Los servicios de salud se benefician de estos modelos en el mejoramiento de los cuidados médicos o la detección de causalidad de enfermedades, pero para su entrenamiento se requieren datos de usuarios que cumplan con los principios de FAIR y CARE. Para ello el autor del presente trabajo entiende la necesidad de una plataforma encargada de recolectar datos cumpliendo los requisitos antes mencionados y ponerlos a disposición de un servicio externo capaz de realizar tareas de MLOps con esta información obtenida de forma consciente. A continuación se describe la arquitectura de referencia TRHEAD para la implementación de plataformas de recolección y uso consciente de datos de sensado móvil y vestible y EvaNotebook como herramienta de MLOps para el despliegue y ejecución de modelos de aprendizaje de máquina.

Capítulo 3. La Arquitectura de Referencia TRHEAD

El objetivo principal de esta arquitectura es ofrecer una plataforma para la recolección y uso de datos que asegure su procedencia y registre las operaciones sobre los mismos manteniendo la trazabilidad. Con el fin de identificar los roles que intervienen en la interacción con la plataforma y las funcionalidades que esta debe ofrecer, se analizaron posibles escenarios para su aplicación.

3.1. Escenarios

Para una mejor comprensión de los casos de uso donde resalta la utilidad de esta arquitectura, se exponen los siguientes escenarios que ejemplifican la interacción de los usuarios con la plataforma. Dicha interacción se divide por roles según las tareas que realizan:

- **Generadores o tenedores de datos:** encargados de subir a la plataforma los datos sanitarios que generan mediante sensado o tienen en su poder (habiéndolos obtenido de forma legítima) y generar automáticamente NFTs que representan los privilegios sobre estos activos, permitiéndoles controlar su acceso y mantener la trazabilidad de su consumo.
- **Curadores de datos:** quienes pueden adquirir estos NFT con crédito interno de la plataforma y con ellos una carta de consentimiento digital para su consumo en la que se especifique el plazo y el alcance de sus privilegios sobre ellos. Esto se hace con el fin de procesar, segmentar y etiquetar los datos para crear un nuevo conjunto de datos con valor añadido que puedan ser de mayor utilidad para los consumidores.
- **Consumidores:** tienen las opciones de consumir los datos en bruto de la plataforma, preferir los datos ya procesados y/o etiquetados por especialistas (curadores de datos), o directamente requerir un modelo de aprendizaje de máquina ya entrenado utilizando los datos que se encuentran disponibles para consumo. Cada una de estas opciones incluye las correspondientes cartas de consentimiento digitales y el pago en crédito de la red a quien puso a disposición los datos; ya sean generadores, tenedores o curadores.

3.1.1. Escenario 1.1 (Generador de Datos):

Descripción corta: Uso de la herramienta en la publicación de datos crudos, consentimiento.

Participantes: Alberto, paciente de pie diabético

Alberto se levanta como cada mañana y antes de realizar cualquier otra actividad toma una foto de la úlcera de pie diabético de la que padece, según la indicación de su médico. Acto seguido sube la foto a la plataforma y esta se comprime antes de subirla a un sistema de almacenamiento en la nube. Las imágenes se guardan en apartados creados por cada estudio clínico para cada usuario, estos apartados ya contienen una carta de consentimiento digital que permite el uso pleno del contenido multimedia que en ellos se almacena. Alberto ha decidido, mediante los estatutos de esta carta de consentimiento que todos los usuarios puedan obtener las capturas de su lesión. Al buzón del paciente llega una notificación de parte de una institución médica que desarrolla una nueva versión de un medicamento para curar la úlcera del pie diabético, explicando que la lesión que se muestra en las fotos subidas al sistema coincide con el tipo de padecimiento que busca aliviar este medicamento experimental. Por lo que hacen una invitación a Alberto a participar de un estudio clínico donde recibirá el compuesto de forma gratuita y en caso de declararse adecuado para tratamiento se recompensará su colaboración con dosis gratuitas durante un año.

3.1.2. Escenario 1.2 (Tenedor de Datos):

Descripción corta: Uso de la herramienta en la publicación de datos crudos, procedencia de los datos

Participantes: Beatriz, neuróloga especialista en pacientes con demencia

Beatriz es una neuróloga especializada en el tratamiento de pacientes con demencia que registra todas sus sesiones como grabaciones de audio. En su clínica privada decidió implementar una promoción al precio de sus consultas que ofrece rebajas a cambio de que los pacientes firmen una carta de consentimiento que permite a Beatriz usar y comercializar libremente sus datos y el material grabado.

Beatriz sube los audios de sus sesiones para el uso libre de los mismos por otros especialistas o investigadores trabajando en el tema u otros afines. Para ello genera cartas de consentimiento digitales

que justifican la procedencia de estos datos. El objetivo de la neuróloga es que aquellos que necesiten evidencia de la procedencia de los datos deban comprar estas cartas de consentimiento, a sabiendas de que este requisito es propio de empresas que desean generar un producto con sus datos. Esto se debe a que para comercializar un servicio de salud, siendo esta una esfera tan crítica para la sociedad, se necesita evidencia de la procedencia de los datos que se usaron en su confección para que este inspire confianza a los pacientes o empresas interesadas. De manera que al comprar las cartas de consentimiento que posee Beatriz, se asegura no solo que los datos son de personas reales, sino que fueron recolectados de manera legal y consciente.

La empresa ModelTrainee está interesada en el conjunto de datos de audio que Beatriz subió a la plataforma con el fin de entrenar un modelo de aprendizaje de máquina para asistir en la detección de demencia en adultos mayores. Para poder comercializar este modelo necesitan registro de la procedencia de los datos pues sus compradores lo solicitan, por lo que a través de la plataforma pagan el monto requerido por la neuróloga para las cartas de consentimiento en moneda de la red. Beatriz recibe dicho monto en su billetera digital menos un porcentaje por costos de transacción.

3.1.3. Escenario 2 (Curador de Datos)

Breve descripción: uso de la herramienta en la publicación de datos procesados, intervención de varios roles, flujo esperado.

Participantes: Centro de Salud MexHealth (tenedor de datos), Carlos (curador de datos), EDUCA (consumidor de datos).

3.1.3.1. Tenedor de datos

El centro MexHealth maneja expedientes médicos de sus pacientes que han sido creados, organizados y actualizados por sus especialistas. Dichas historias clínicas contienen datos de ritmo cardíaco de pacientes con diferentes enfermedades durante su período de sueño. Los datos contenidos en estos registros pertenecen a la institución, ya que sus pacientes han firmado una carta de consentimiento para su uso, motivados por una rebaja en los servicios que reciben de la institución. Estas historias clínicas se cargan

en la plataforma junto con las cartas de consentimiento de los pacientes para garantizar la fidelidad de su procedencia. El centro permite que estos datos estén disponibles de forma gratuita, pero no las cartas de consentimiento que se suben en otro apartado del almacenamiento en la nube. Para poder acceder a ellas, el consumidor debe comprometerse a dar participación gratuita a la institución de cualquier resultado obtenido en los siguientes términos:

- Si el resultado es un servicio: licencia permanente para el uso de dicho servicio y versiones posteriores.
- Si se trata de un conjunto de datos tratados: Acceso para el consumo de los mismos y versiones posteriores.
- Si se trata de un artículo de valor científico: Mención de la institución como fuente de los datos.

3.1.3.2. Consumidor de datos

El centro de investigación y educación superior EDUCA está llevando a cabo un estudio médico para demostrar la relación entre el cáncer de mama y el insomnio en mujeres. Para ello, la institución requiere datos de frecuencia cardíaca de mujeres con cáncer de mama en diferentes fases de la enfermedad. Estos se usarán para entrenar un modelo de aprendizaje de máquina que automatice la detección de patrones irregulares de sueño que puedan estar relacionados con cáncer de mama. Los datos deben estar libres de registros erróneos debido a fallos del sensor o registros vacíos debido a que las pacientes se han quitado el sensor. También deben estar etiquetados por fase del cáncer y agrupados por conjuntos de edades. Desafortunadamente, los datos de ritmo cardíaco proporcionados por MexHealth no cumplen con estos requisitos, por lo que EDUCA publica dentro de la plataforma una convocatoria para curadores de datos que puedan encargarse de esta tarea.

3.1.3.3. Curador de datos

Carlos es un curador de datos con experiencia en el análisis de ritmo cardíaco y decide aceptar la tarea de procesar los archivos de datos de acuerdo a las necesidades del centro educativo. Sube los resultados

de su trabajo a la plataforma poniéndolos a disposición para su consumo bajo la condición: Abono en su billetera electrónica de la cantidad equivalente a 500 MX en BTC; y el centro MexHealth queda exento de pago de acuerdo con los términos que estipula el acceso a las cartas de consentimiento de los pacientes.

3.1.3.4. Resultado

EDUCA se interesa por el trabajo de Carlos gracias a que está respaldado por cartas de consentimiento de pacientes, por lo que son datos válidos y legalmente obtenidos. Luego procede a comprar el conjunto de datos por la cantidad requerida. Con el fin de comercializar su modelo de aprendizaje automático con evidencia de la veracidad de los datos utilizados, acceden a las cartas de consentimiento de MexHealth a cambio de una licencia de uso permanente del modelo para dicha institución sanitaria.

3.1.4. Escenario 3 (Consumidor de datos)

Breve descripción: uso de la herramienta en el consumo de datos en bruto, no refutabilidad.

Participantes: Daysi, investigadora de una empresa farmacéutica.

Daysi investiga el efecto de una crema hidratante en mujeres jóvenes de piel clara para la empresa farmacéutica Vitalia. Para ello, recopila imágenes de rostros de mujeres de piel clara de entre 15 y 30 años. En la interfaz del consumidor de la plataforma, hace clic en el botón “Nuevo estudio” e introduce los criterios de discriminación de los participantes antes de confirmar la creación del estudio. A continuación, el sistema envía una notificación a un grupo de usuarias cuyos perfiles coinciden con los criterios, dándoles la opción de participar en el estudio. Las jóvenes deben subir imágenes de su rostro cada día antes de aplicarse la crema y después. Las imágenes se suben en secciones creadas para el estudio para cada usuaria, las cuales ya contienen una carta de consentimiento que permite el uso completo del contenido multimedia almacenado en ellas. Con la recopilación obtenida y tras su investigación, Daysi consigue demostrar que la crema reduce la aparición de manchas en la piel. Los resultados de la investigación se publican en un artículo junto con algunas imágenes de las jóvenes para visualizar más fácilmente el efecto de la crema.

El departamento jurídico de la empresa farmacéutica notifica a Daysi que una de las participantes cuya foto fue publicada alega que esto se hizo sin su consentimiento y le ha causado daños psicológicos por lo que planea demandar a la institución. La investigadora procede a generar una versión digital de la carta de consentimiento firmada por la denunciante a través de la plataforma y la envía al departamento jurídico para su disposición. El asunto se resuelve sin escalar a mayores.

3.2. Funcionalidades necesarias en una plataforma de TRHEAD

Como se puede observar en los escenarios expuestos, para suplir las necesidades de los roles involucrados en el flujo de datos de esta arquitectura 3 se necesitan asegurar un conjunto de funcionalidades:

- Una interfaz que permita a los usuarios crear, cargar y consumir datos de salud.
- Interacción con los sensores de dispositivos móviles y vestibles para obtener los datos que estos generan.
- Un registro confiable de las operaciones que tienen lugar en la plataforma para asegurar la trazabilidad de los datos.
- Un sistema para el almacenamiento de los datos de sensado móvil y vestible, además de sus correspondientes cartas de consentimiento.
- Un servicio de aprendizaje automático que pueda obtener los datos de salud conscientemente/legalmente recopilados para el entrenamiento de modelos de aprendizaje automático de procedencia verificada.

Estas funcionalidades deben implementarse respetando la privacidad y el consenso del usuario, puesto que una de sus finalidades es el entrenamiento de modelos de Aprendizaje de Máquina utilizando estos datos obtenidos de forma consciente. Para cumplir con estos requisitos, la arquitectura debe contar con los componentes que permitan la realización de las tareas en las distintas etapas del flujo de datos, los cuales se describen a continuación.

3.3. Componentes de TRHEAD

TRHEAD se compone de una DApp (Aplicación Descentralizada) como interfaz gráfica para la interacción del usuario, encargada de ejecutar las librerías que permiten la comunicación con la Blockchain. La propia “cadena de bloques” funciona como un registro descentralizado, inmutable y público de las operaciones realizadas en la plataforma y de los NFTs (contratos inteligentes) que representan los derechos sobre los datos. Otro componente es el sistema de almacenamiento en la nube, que se utiliza para almacenar tanto los datos sanitarios como los metadatos asociados a los NFT que se registran en la cadena. Con el fin de consumir los datos almacenados en la nube para el entrenamiento consciente de modelos de aprendizaje de máquina, se propone el uso de una herramienta externa capaz de obtener datos a través de protocolos de Internet para su procesamiento como lo son Google Colab, Amazon Sagemaker, etc. Los modelos entrenados se distribuirán como instancias autónomas de un modelo maestro, que contendrá las cartas de consentimiento digitales que autorizan el uso de los datos de entrenamiento. Para una mejor comprensión de los elementos de la estructura y la interacción entre los mismos refiérase a la Figura 1.

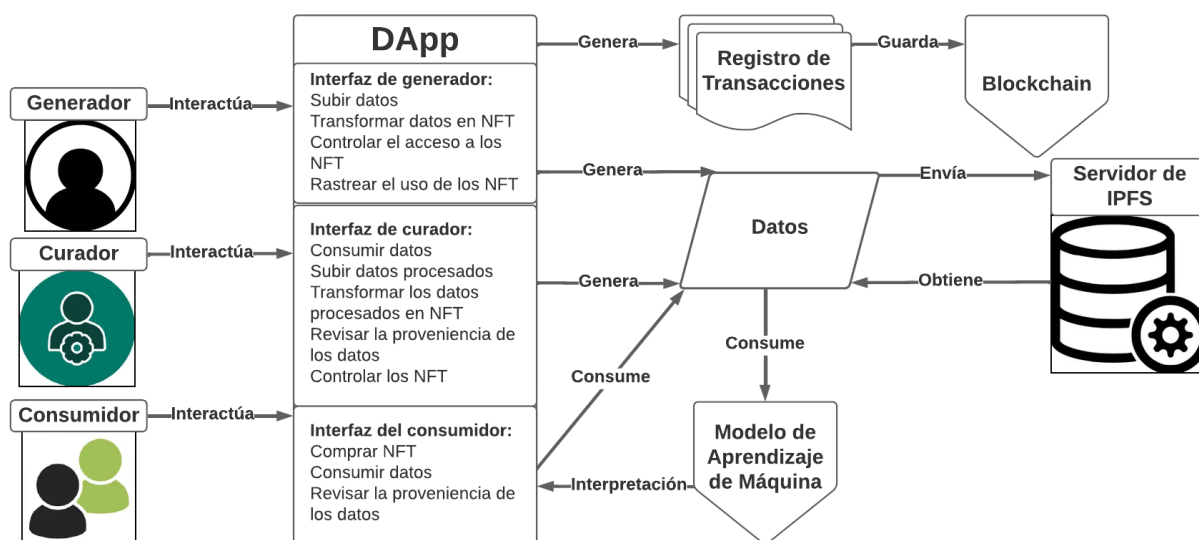


Figura 1. Estructura de la arquitectura de referencia TRHEAD

3.4. DApp

La Aplicación Descentralizada funciona como interfaz gráfica con la que interactúa el usuario por lo que debe ser intuitiva, de manera que los usuarios no técnicos puedan aprender a usarla. Su función principal es la de comunicarse con la Blockchain para realizar las operaciones necesarias en la gestión de los datos de sensado. Para llevar a cabo sus funciones necesita acceso a la clave privada de la billetera electrónica (e-wallet) del usuario, la cual se usará para firmar las transacciones que este realice dentro de la plataforma que se construya usando esta arquitectura. La DApp es también la encargada de recopilar o sensar los datos de salud del paciente, así como mostrar aquellos que ya están disponibles para su comercialización a los consumidores.

3.5. Sistema de almacenamiento en la nube

Tanto la información sanitaria de los pacientes como los metadatos asociados a los NFT que los representan en la Blockchain, deben ser almacenados de forma que se encuentren disponibles en todo momento. Se debe considerar también que la cantidad de datos a almacenar es extensa e irá en aumento a medida que se use la plataforma, por lo que sería riesgoso depender de un solo servicio de almacenamiento. Permitir a una sola institución el acaparar semejante cantidad de información podría resultar en notables fugas de información y otras formas de manipulación incorrecta de la misma. Proponemos que el almacenamiento debe hacerse en la nube para permitir su acceso desde cualquier parte del mundo con una conexión a Internet y el uso de un servicio descentralizado, donde sean los propios productores de datos quienes se responsabilicen de su integridad y disponibilidad.

3.6. Blockchain

A diferencia de un servidor centralizado, las decisiones en una “cadena de bloques” se llevan a cabo mediante mecanismos de consenso en los que participan directamente sus nodos. Esta naturaleza descentralizada protege a la red de ataques de adversarios, ya que el consenso debe alcanzarse por mayoría y, por tanto, cuantos más nodos haya, mayor será la seguridad. Los contratos inteligentes son programas que se ejecutan en la propia red y se vuelven inmutables una vez desplegados en ella, sirviendo como

ejecutores y controladores de las operaciones que en ella se realizan (Hassanein et al., 2022). En la arquitectura TRHEAD, la “cadena de bloques” es el ente regulador de todas las operaciones, almacena los registros de uso de datos y controla los privilegios que se tienen sobre los mismos. Siempre que se realiza una operación sobre los datos queda registrada la dirección de e-wallet del iniciador de la misma y de todo aquel que intervino o se vio envuelto en ella, lo que permite una trazabilidad de la misma de forma automática, pública e inmutable.

Los contratos inteligentes son los encargados de llevar a cabo estas operaciones cuando se cumplen las condiciones para su ejecución y su principal ventaja radica en que automatizan el acuerdo entre pares, prescindiendo de la intervención de un tercero. Esto permite que la cadena se convierta en una especie de libro de cuentas el cual todos pueden consultar, pero solo lo editan aquellos que superan una prueba establecida por la red de nodos para asegurar que no se introduzcan registros inválidos. A esta prueba se le conoce como mecanismo de consenso y generalmente está enfocada a evitar que un grupo de individuos malintencionados manipulen la red en su beneficio, ya sea centralizando las operaciones en sus nodos o realizando ataques adversarios.

3.7. Modelo de Aprendizaje de Máquina

El aprendizaje automático representa un cambio de paradigma en el descubrimiento de la causalidad de las afecciones a la salud y el mejoramiento en general de los sistemas sanitarios. La ciencia de datos aplicada a la medicina permite que los registros médicos se conviertan en materia prima para entrenar modelos que ayudan a los profesionales de la salud a precisar diagnósticos y a los pacientes a ser más pro-activos hacia el seguimiento y control de su enfermedad. Estos permiten el uso de varios tipos de datos; dígase imágenes, audio, series de tiempo, etc; y al automatizar su tratamiento evitan el sesgo y arrojan mejores resultados que los métodos tradicionales donde intervienen los investigadores.

En la arquitectura propuesta estos modelos son consumidores finales para datos éticamente recopilados y serán distribuidos como versiones básicas enfocadas en funciones específicas como detección de somnolencia, emociones o patrones de conducta típicos de una enfermedad. Su objetivo final es el de ser usados para obtener interpretaciones a partir de un conjunto de datos o ser re-entrenados mediante el proceso de “transfer learning” (Weiss et al., 2016). Se necesita un entorno con poder computacional suficiente para la ejecución de los algoritmos de alto consumo de recursos con los cuales se construyen estos

modelos y esto implicaría depender de equipos de alto desempeño que suelen ser en extremo costosos económica y ambientalmente. Esta forma centralizada de procesamiento está sujeta a la posibilidades de desperfectos técnicos en los equipos o intervención malintencionada en el modelo resultante por parte de aquellos con acceso al mismo.

Un enfoque descentralizado permitiría distribuir la carga de cómputo entre varios equipos mediante servicios de división, distribución y calendarización de tareas como Amazon SageMaker, reduciendo el tiempo de entrenamiento de los modelos mediante la filosofía "divide y vencerás". Sin embargo el hecho de usar un servicio de una compañía es una forma de centralizar las operaciones e implicaría que esta tendría acceso a los datos de entrenamiento, además los consumidores de los modelos estarían a merced de las disposiciones comerciales que impongan las reglas de mercado.

Una alternativa ideal sería un servicio capaz de distribuir el entrenamiento entre varios nodos de una "cadena de bloques" mediante el uso de una librería como Tensor Flow Distribuido. El mecanismo de consenso de la cadena se usaría para elegir y recompensar a los nodos que participen en las tareas según su confiabilidad y la magnitud de su aporte al resultado final. Los datos almacenados en el sistema de almacenamiento distribuido serían recuperados por un contrato inteligente corriendo en la Blockchain y sus direcciones de almacenamiento puestas a disposición del servicio de entrenamiento de modelos distribuido para la creación de modelos que se almacenarían de vuelta y se registrarían en la cadena como nuevos activos.

3.8. Flujo de datos entre los componentes

Para ejemplificar la interacción entre los componentes de la arquitectura de referencia THREAD presentamos un diagrama de secuencia 2 que muestra cómo un sistema usaría los componentes propuestos en dicha arquitectura para implementar el Escenario 1.1 descrito en la sección 3.1.1.

Una vez descrita la arquitectura de referencia con sus principales componentes y funcionalidades, procedemos a describir las implementaciones, prueba y campaña de sensado de sistemas que cumplen con los principios de TRHEAD.

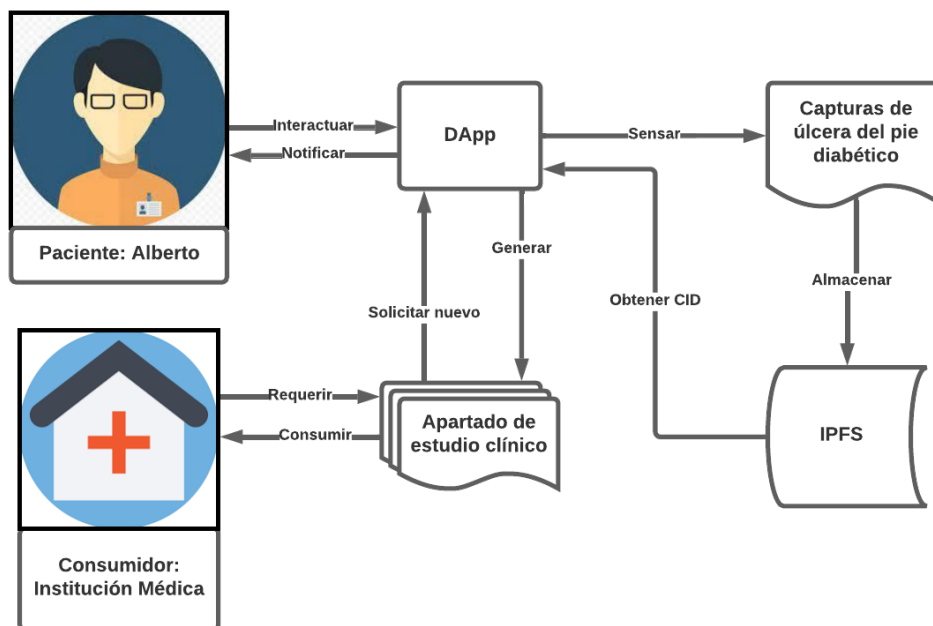


Figura 2. Flujo de datos entre los componentes del Escenario 1.1

Capítulo 4. Implementaciones de la arquitectura de referencia THREAD para el sensado móvil

Con el fin de poner en práctica la arquitectura de referencia planteada se realizó una implementación de la misma en la creación de un sistema para el sensado, recopilación y consumo de datos. La motivación principal de esta implementación fue la necesidad de conocer la viabilidad de un sistema basado en TRHEAD para su despliegue en escenarios reales. Para cumplir con la función esperada, la plataforma debe permitir recopilar datos del usuario a través de los sensores del dispositivo, subirlos a un sistema de almacenamiento distribuido y servirlos a través de una API Rest (Ranga & Soni, 2019) para su consumo. Para comprobar el funcionamiento correcto de la API se implementó un programa capaz de obtener los datos que esta ofrece y crear con ellos un “dataset”, estructura que comúnmente se usa en el entrenamiento de modelos de ML. Todas las operaciones antes mencionadas se registran en un contrato inteligente de la Blockchain. Los componentes del sistema y las tecnologías usadas en su desarrollo se exponen a continuación.

4.1. Componentes

La plataforma está compuesta por una interfaz gráfica web, un servidor que implementa el API Rest, un Sistema de Gestión de Base de Datos (SGBD), un contrato inteligente y el programa que consume la API. Intervienen de forma menos directa otros componentes como el sistema de almacenamiento distribuido y la “cadena de bloques” donde corre el contrato inteligente.

4.1.1. Interfaz gráfica

La interfaz consiste de una Aplicación Web Progresiva (PWA) (Bjørn-Hansen et al., 2017) codificada usando la librería VueJS del lenguaje Javascript. Es la herramienta mediante la cual el usuario interactúa con el sistema para crear, subir y controlar sus activos digitales (datos de salud), así como para consumirlos de otros usuarios. Cada una de estas operaciones cuenta con una interfaz propia y responden a la autenticación del usuario, de manera que solo aquellos registrados pueden subir o eliminar activos.

4.1.1.1. Subir datos de salud

La vista para subir datos de salud permite cargar archivos locales que el paciente haya obtenido de su doctor o sensado por su cuenta y subirlos al almacenamiento distribuido como se muestra en la Figura 3. Para ello el usuario debe estar autenticado y ofrecer etiquetas que describan la naturaleza de los datos subidos, para facilitar su clasificación y posterior consumo. Finalmente los firma con sus credenciales y se convierten en activos de la plataforma que podrá encontrar en el apartado correspondiente a su usuario del listado de activos.

Utilizar una tecnología web permite que la interfaz funcione tanto en dispositivos móviles como en portátiles y escritorio, de manera que el campo de selección de archivo lanza el explorador de archivos independientemente del sistema operativo anfitrión. Los archivos que ha subido el paciente/usuario se listan con las correspondientes opciones de descargarlos o eliminarlos. Los datos del paciente pueden ser introducidos o dejar vacíos los campos para un envío anónimo de los mismos. Para etiquetar los archivos que se desean compartir existen etiquetas predefinidas en el campo seleccionable o el usuario puede agregar aquellas competentes a la naturaleza de sus datos. Finalmente firma su nuevo activo con su clave única de usuario y este se muestra en el listado global con sus correspondientes propiedades; como autor, precio y naturaleza de los datos.

Figura 3. Interfaz de gestión de archivos

4.1.1.2. Sensado de datos de acelerometría

La vista para el sensado de datos de acelerometría, apreciable en la Figura 4, hace uso de las bibliotecas “Gyroscope” y “Accelerometer” de Javascript, disponibles en el navegador, lo que permite acceder al giroscopio y acelerómetro del dispositivo independientemente del sistema operativo. Estos son sensores que miden la aceleración lineal y aceleración angular respectivamente en un dispositivo. Por ejemplo, la reciente generación de teléfonos inteligentes incluye por defecto sensores de acelerómetro basados en MEMS (Sensores Microeléctromecánicos). Su acelerómetro mide fuerzas de aceleración constantes (gravedad), variables en el tiempo (vibraciones) y cuasiestáticas (inclinación), que afectan al dispositivo en los tres ejes (x, y, z) en metros por segundo al cuadrado (Grouios et al., 2023).

En esta vista se registran los valores de aceleración y ubicación, además de una marca del tiempo que transcurre durante este proceso. Los botones de “Iniciar”, “Detener” y “Reiniciar” controlan el registro de los datos que se obtienen del sensor, mientras que el botón “Guardar” se usa para descargar un dataset en forma de archivo CSV (Coma Separated Values) con el contenido registrado.



Figura 4. Interfaz de captura de datos de acelerometría

4.1.1.3. Listado de activos

La vista para el listado de activos que se puede observar en la Figura 5 presenta todos los datos médicos en forma de contratos firmados y agrupados por su generador, tenedor o curador, en dependencia del rol del usuario que los subió a la plataforma. Se mostrará el nombre, tamaño, formato, precio y etiquetas de los archivos siempre que estos estén disponibles y de forma opcional el nombre del paciente, teniendo en cuenta que los datos pueden compartirse de manera anónima.

El botón “Comprar” permite a un curador o consumidor obtener estos activos con la correspondiente carta de consentimiento digital que autoriza su uso, siempre que se abone la cantidad especificada para el mismo, la cual es cero en caso de los datos de salud que se aportan para dominio público. De contar con el crédito suficiente en moneda de la red (Blockchain), se descuenta el mismo de su billetera electrónica y se transfiere a la del propietario de los activos restando un impuesto por transacción. Acto seguido se descargan los datos en el equipo del comprador y queda registrada dicha operación, de manera que exista un registro probatorio de la adquisición de estos datos de forma legal para su uso consciente.

The screenshot displays a user interface for viewing active contracts. At the top, there is a toggle for 'Expandir' and a title 'CONTRATOS ACTIVOS POR USUARIO'. Below this, the user 'GRENETER CORDOVES DELGADO' is identified as the 'Curador de datos'. The interface lists two data contracts, each with a 'COMPRAR' button.

INFORMACIÓN DE LOS DATOS:			COMPRAR
NOMBRE DE ARCHIVO: GRENETER-20230209T021240Z-001.	TAMAÑO DE ARCHIVO: 3360500 B	FORMATO DE ARCHIVO: APPLICATION/ZIP	
ETIQUETAS: Grabación de voz, Interacción con el robot EVA, Estados de ánimo			
PRECIO: 0	PACIENTE: GRENETER CORDOVES DELGADO		
INFORMACIÓN DE LOS DATOS:			COMPRAR
NOMBRE DE ARCHIVO: FEATURES.RAR	TAMAÑO DE ARCHIVO: 430798 B	FORMATO DE ARCHIVO:	
ETIQUETAS: Vector de características de emociones			
PRECIO: 50	PACIENTE: GRENETER CORDOVES DELGADO		

Below the list, two more users are visible as 'Curador de datos': ELIAF YAHIR GARCÍA LOYA and CHRISTIAN GONZÁLEZ GARCÍA.

Figura 5. Listado de activos en la plataforma

4.1.2. Servidor Web

El servidor web, implementado en ExpressJS, se encarga de la compilación y despliegue de nuevos contratos inteligentes, la gestión de usuarios del sistema y las peticiones al servicio de almacenamiento. El uso de la librería “solc” de Javascript permite que se compilen nuevos contratos inteligentes usando un archivo de Interfaz de Aplicación Binaria (ABI) para el lenguaje de programación Solidity. La gestión de usuarios incluye el manejo de su información personal, las direcciones de los contratos que han desplegado y los datos de salud que han subido a la plataforma. Mediante funciones escritas en el lenguaje Javascript y la interacción con el SGBD se podía gestionar el CRUD (Crear, Obtener, Actualizar, Borrar) de usuarios mediante llamadas de tipo HTTP al servidor, gracias a que fue implementado bajo el patrón arquitectónico REST (Feng et al., 2009).

Como Sistema Gestor de Base de Datos se eligió MongoDB, usando su versión en la nube conocida como Atlas. MongoDB es una tecnología NoSQL destinada al almacenamiento de documentos y se caracteriza por su flexibilidad y escalabilidad. Esto permitió crear estructuras de código para el almacenamiento de la información de los usuarios y los datos que estos suben a la plataforma con la posibilidad de adaptarse a cambios en el modelo de negocio. De esta forma en caso de ser necesario se podrían agregar nuevos nodos al “clúster” de almacenamiento para ampliar su capacidad.

Se implementó además una API para el consumo vía web de los datos etiquetados como “aportación gratuita”, para su consumo por una herramienta externa.

4.1.3. Contrato Inteligente y la Blockchain

A la hora de elegir un lenguaje para la implementación del contrato inteligente se tuvo en cuenta la curva de aprendizaje y la compatibilidad de este con el amplio ecosistema de Blockchains existentes en la actualidad. La red Ethereum fue la primera plataforma para contratos inteligentes y ha servido como base para la creación de nuevas plataformas con la posibilidad de desplegar estos programas. La red Polkadot expande esta filosofía a un enfoque de compatibilidad e interoperabilidad, mediante la creación de subcadenas (parachains) que desprenden de este tronco común pero se implementan con nuevas características que permiten ampliar el abanico de posibilidades de uso.

Una de estas “parachains” es Moonbeam, que además de poseer esta compatibilidad e interoperabilidad, también permite la comunicación con la cadena Ethereum y ofrece un ambiente de implementación y despliegue cómodo para desarrolladores. Habiendo elegido la plataforma se implementó el contrato en el lenguaje de programación Solidity con las funcionalidades necesarias para desplegarlo, detenerlo, cambiar su propietario y removerlo de la red. Se agregaron además variables para almacenar datos de la carta de consentimiento de uso de datos como el costo, expiración del permiso, dirección de almacenamiento y restricciones de uso, las cuales se pueden apreciar en la Figura 6.

El contrato implementa las funciones “pausable, burnable y ownable” que permiten detener la ejecución, destruir y cambiar la propiedad del contrato respectivamente. Esto permite que sea el paciente quien decida por cuánto tiempo estarán disponibles sus datos y a quién ceder sus privilegios sobre ellos. De igual manera un curador de datos podría detener la distribución de uno de los datasets que confeccionó al descubrir un error de anotación. La función “mint” envía la información del contrato a la cadena para que esta sea clonada entre los nodos de la misma y comience su ciclo de ejecución ininterrumpida hasta que su propietario encuentre pertinente.

```
// SPDX-License-Identifier: MIT
pragma solidity ^0.8.9;

import "@openzeppelin/contracts@4.9.3/token/ERC20/ERC20.sol";
import "@openzeppelin/contracts@4.9.3/token/ERC20/extensions/ERC20Burnable.sol";
import "@openzeppelin/contracts@4.9.3/security/Pausable.sol";
import "@openzeppelin/contracts@4.9.3/access/Ownable.sol";

contract ConsentLetter is ERC20, ERC20Burnable, Pausable, Ownable {
    uint public state;
    uint public payment;
    uint public expiration;
    string public storageURI;
    string public usageRestrictions;
    constructor() ERC20("ConsentLetter", "CONL") { state = 0; }

    function pause() public onlyOwner {
        _pause();
    }

    function unpause() public onlyOwner {
        _unpause();
    }

    function mint(address to, uint256 amount) public onlyOwner {
        _mint(to, amount);
    }

    function _beforeTokenTransfer(address from, address to, uint256 amount)
        internal
        whenNotPaused
        override
    {
        super._beforeTokenTransfer(from, to, amount);
    }
}
```

Figura 6. Código del contrato inteligente “Consent Letter” escrito en el lenguaje Solidity

4.1.4. Aplicación externa

Para comprobar el correcto funcionamiento de la API que provee los datos etiquetados como “aportación gratuita”, se creó una aplicación de Django capaz de obtenerlos vía web y crear con ellos un “dataset”. A pesar de que no se entrenó un modelo, esta aplicación simple basta para comprobar que otras herramientas como Google Colab podrían consumir estos datos en la realización de tareas de aprendizaje automático. La API consulta los datos médicos de la base de datos en Atlas y filtra aquellos con la etiqueta especificada, estos llevan registrado el CID que los identifica en el servidor IPFS donde fueron almacenados. La respuesta de la API a una petición HTTP/GET es una lista de los CIDs de los datos filtrados. Este flujo está representado en la Figura 7 para una mejor comprensión.

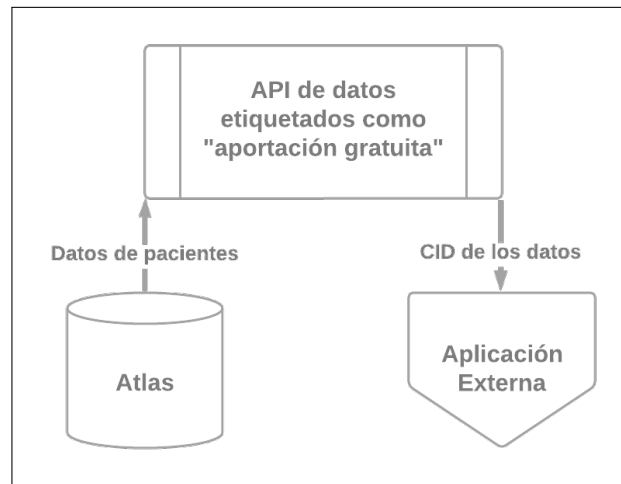


Figura 7. Flujo de las direcciones de los datos para su consumo final

4.2. Pruebas

Con el fin de comprobar la viabilidad de la plataforma en un entorno real se realizó una prueba con trece estudiantes de la asignatura de Ciencia de Datos para Sensores del Posgrado de Ciencias de la Computación del CICESE (Centro de Investigación Científica y Educación Superior de Ensenada). La prueba consistía en realizar la captura, segmentación y etiquetado de datos como parte de un proyecto de clase para detección de actividades usando sensores inerciales. El motivo de elección de esta muestra fue el contexto con el que cuentan los sujetos por encontrarse cursando la asignatura de Ciencia de Datos para Sensores, donde conocen la importancia de la correcta manipulación y etiquetado de los datos.

4.2.1. Procedimiento

Cada uno de los sujetos debía entrenar un modelo de aprendizaje de máquina a partir de las características obtenidas de un conjunto de datos de acelerometría. Estos datos debían obtenerlos a partir de varias sesiones de sensado usando el acelerómetro de su dispositivo móvil durante actividades como correr, saltar y caminar. Para el entrenamiento del modelo contaban con un conjunto de datos base, pero se requería que este clasificara correctamente un conjunto de entradas de prueba definidas por su profesor que difería de los datos usados en el entrenamiento. Esto forzaba a los estudiantes a intercambiar entre ellos el resultado del sensado o incluso el conjunto de características extraídas de este para robustecer el modelo, por lo que la plataforma se propuso como vía para este proceso. La prueba duró una semana y durante este período el flujo de información dentro de la plataforma superó los tres Giga-bytes lo cual resulta positivo teniendo en cuenta que el peso de los datos manejados no superaba el orden de los Kilo-bytes.

4.2.2. Encuestas de experiencia de usuario

Con el fin de evaluar la experiencia del usuario al utilizar la plataforma se confeccionó un cuestionario donde debían responder preguntas sobre su interacción con la misma. El cuestionario se compone de veintinueve preguntas divididas en tres secciones correspondientes a Medición de Aceptación de la Tecnología (TAM), Conformidad con la tecnología e Intención de Uso; con diecinueve, doce y tres preguntas respectivamente. Al final de cada sección se incluyó un espacio para comentarios donde el encuestado podía expresar cualquier inquietud no cubierta en el cuestionario. Las preguntas en cada sección también se dividieron por los tres roles principales de la plataforma, los cuales se les presentaron a los encuestados en forma de escenario de la siguiente forma:

- **Generador/Tenedor de datos:** “Para la siguiente sección ubíquese en el rol de Carlos, quien es un curador de datos con experiencia en audio de voz y se da a la tarea de procesar los datos de los expedientes médicos de pacientes con sospecha de demencia. Debe procesar los audios de manera que no exista ruido exterior, silencios, interrupciones y se escuche claramente la voz del paciente. Además debe etiquetar los audios procesados con los datos del paciente. Se deben cumplir dichos requisitos para poder publicarlos al banco de datos DementiaBank para que se pongan a disposición de uso. Elige subirlos a la plataforma de Creación y Distribución de Datos de Sensado para que

sirva de almacenamiento y sistema de gestión.”

- **Curador de datos:** “Para la siguiente sección ubíquese en el rol de un paciente con úlcera del pie diabético que es usuario de la plataforma y se llama Alberto. Como parte de su rutina diaria se levanta cada mañana y antes de realizar cualquier otra actividad toma una foto de la úlcera de pie diabético de la que padece, según la indicación de su médico. Acto seguido sube la foto a la plataforma y esta se comprime antes de subirla a un sistema de almacenamiento en la nube. Las imágenes se guardan en apartados creados por cada estudio clínico para cada usuario, estos apartados ya contienen una carta de consentimiento digital que permite el uso pleno del contenido multimedia que en ellos se almacena. Alberto ha decidido, mediante los estatutos de esta carta de consentimiento que todos los usuarios puedan consumir las capturas de su lesión que en ella se almacenan.”
- **Consumidor de datos:** “Para la siguiente sección ubíquese en el rol de Daysi, quien investiga el efecto de una crema hidratante en mujeres jóvenes de tez blanca para la farmacéutica Vitalia. Para ello recopila imágenes del rostro de mujeres de tez blanca entre 15 y 30 años. En la interfaz de consumidor de la plataforma presiona el botón “Nuevo estudio” e introduce los criterios de discriminación de los participantes según lo antes mencionado antes de confirmar la creación del estudio. Acto seguido el sistema envía una notificación a un grupo de usuarios cuyos perfiles coinciden con los criterios con la opción de participar del estudio. Las jóvenes deben subir imágenes de su rostro cada día antes de aplicar la crema y posterior a esto. Las imágenes se suben en apartados creados para el estudio por cada usuario, los cuales ya contienen una carta de consentimiento que permite el uso pleno del contenido multimedia que en ellos se almacena. Con la recopilación obtenida y tras su investigación, Daysi logra comprobar que la crema reduce la aparición de manchas en la piel. Los resultados de la investigación son plasmados en un artículo y publicados junto a algunas imágenes de las jóvenes para visualizar con más facilidad el efecto de la crema.”

4.2.3. Resultados de la encuesta

4.2.3.1. Cuestionario de intención de uso

A la pregunta: ¿Qué tan probable es que use una herramienta con estas funcionalidades para las tareas indicadas en el rol anterior por sobre las formas tradicionales de hacerlo como Google Drive, FileCoin,

DropBox, Firefox Send, etc?; los encuestados debían responder con una sola de las opciones dadas: “Altamente probable”, “Probable”, “Poco probable” o “No la usaría en absoluto”. Esta se realizó una vez por cada escenario planteado y los resultados arrojados se pueden observar en la Figura 8.

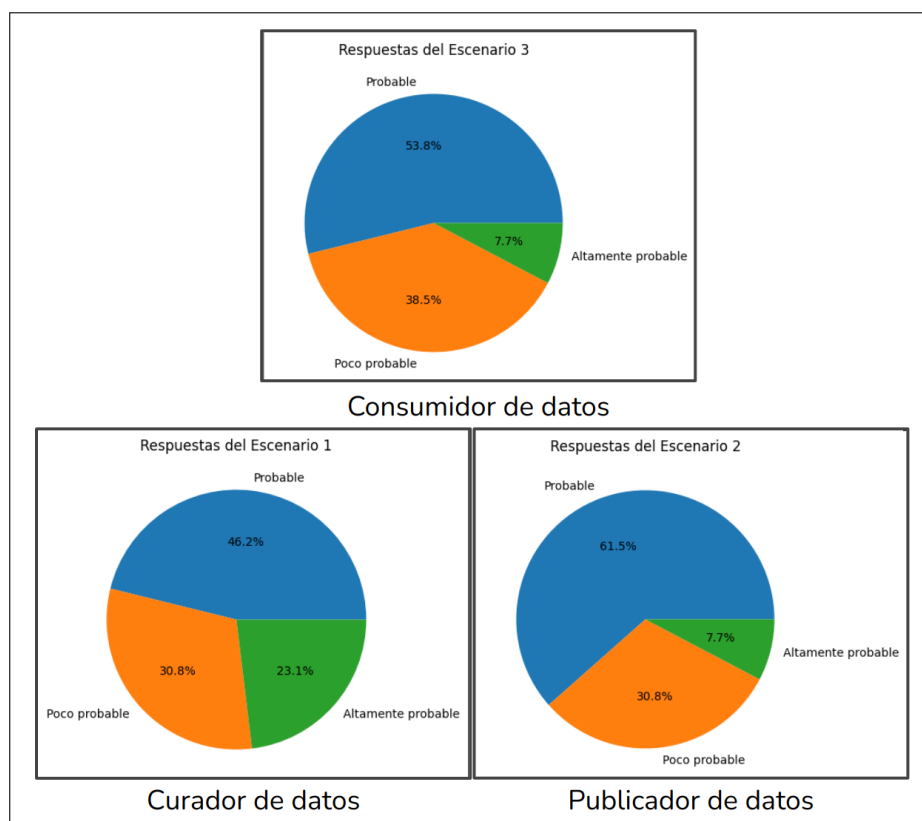


Figura 8. Resultados de los cuestionarios de Intención de Uso

A pesar de que en la mayoría de los escenarios las respuestas fueron mayormente positivas parece haber una inclinación más clara en la intención de uso desde el rol de Curador de Datos. Esto podría deberse a la vocación de los integrantes de la muestra teniendo en cuenta su contexto como estudiantes del curso de Ciencia de Datos para Sensores.

4.2.3.2. Cuestionario de aceptación de la tecnología (TAM)

Para el TAM (Marangunić & Granić, 2015) se solicitó a los encuestados que, partiendo de la experiencia que tuvieron en la interacción con el sistema, evaluaran la utilidad y usabilidad de la plataforma a partir de un conjunto de enunciados; de los cuales debían indicar qué tan satisfechos estaban en una escala Likert de 1 a 5 (donde 1 es nada satisfecho y 5 completamente satisfecho). Esta sección de 12 preguntas

se dividió en seis preguntas relacionadas con la utilidad y las restantes con la usabilidad de la plataforma.

Para la sección de utilidad se obtuvo un promedio de respuestas de 3.6 y una moda de tres, lo cual implica cierta ambigüedad en las respuestas. Esto se puede apreciar con mayor facilidad en el Figura 9 y podría deberse a que algunos de los participantes obtuvieron los datos directamente de sus compañeros.

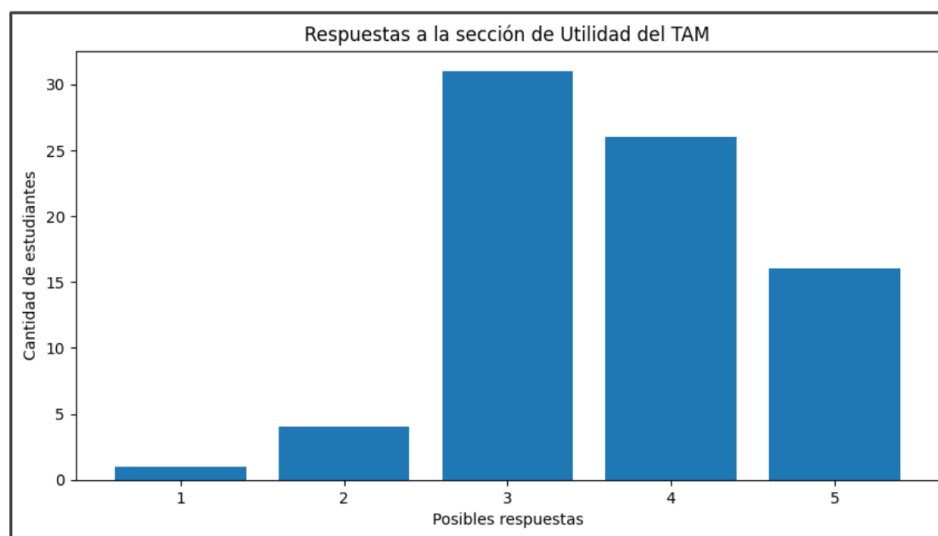


Figura 9. Resultados de los cuestionarios de TAM, sección de utilidad

En la sección de usabilidad los resultados son más positivos con un promedio de 3.9 y una moda de cuatro, indicando que las opiniones de los encuestados perciben la tecnología propuesta como fácil de usar. Esto se puede observar en la Figura 10.

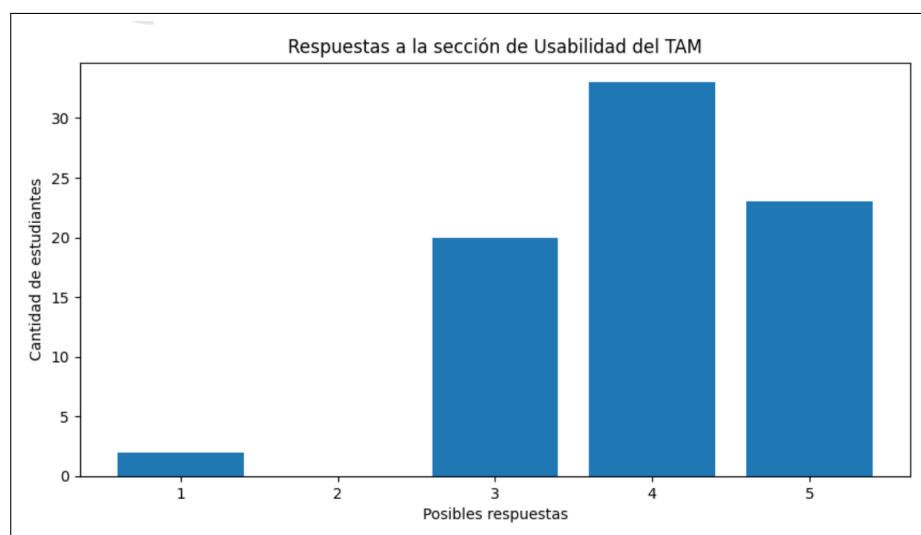


Figura 10. Resultados de los cuestionarios de TAM, sección de usabilidad

Se muestra cierta ambigüedad en las respuestas a las preguntas de utilidad percibida que disminuye en las correspondientes a usabilidad. Podríamos concluir de manera general que la aplicación se percibió más fácil de usar que útil, esto podría ser consecuencia de haberse utilizado en una tarea corta y la falta de interoperabilidad con herramientas de aprendizaje automático como Google Colab, la cual usaron la mayoría de los encuestados en la fase final de su tarea.

4.2.4. Cuestionario de conformidad con la tecnología

Para esta sección se presentaron un conjunto de afirmaciones a los encuestados con el fin de que respondieran qué tan de acuerdo estaban con las mismas. Para ello se les pidió que partieran del punto de vista ofrecido en cada uno de los escenarios planteados en el cuestionario. Las opciones de respuesta se dieron en escala Likert, donde el encuestado debía seleccionar una de las siguientes: “Totalmente en desacuerdo”, “En desacuerdo”, “Ni de acuerdo ni en desacuerdo”, “De acuerdo” y “Totalmente de acuerdo”. Los resultados, expuestos en la Tabla 1, mostraron una aceptación generalizada por lo que se consideran validados los escenarios pero es notable que uno de los individuos mantenía una postura de desacuerdo en todas las propuestas relacionadas con la trazabilidad de los datos al paciente origen.

En las secciones de comentarios proyectaba esta inquietud con enunciados como “Considero que compartir los datos no da derecho a invadir la privacidad de los demás”, demostrando un interés genuino en la protección de la privacidad de los usuarios. Esta predisposición puede estar condicionada por los métodos tradicionales para asegurar la **procedencia** de los datos que suelen ser invasivos para la privacidad del paciente. En otra ocasión comenta que *“Como consumidor de datos debo asegurarme de la confiabilidad y veracidad de los mismos, y ser capaz de seleccionar filtros para los datos que deseo”*, de lo cual se puede inferir que comprende la necesidad de contar con **“data provenance”** pero atribuye al consumidor el deber de asegurarla sin comprometer la privacidad del paciente.

4.3. Conclusiones parciales del experimento

Los datos recolectados de la encuesta arrojan resultados muy positivos en general, sin embargo por la ambigüedad mostrada en el TAM comprendimos la necesidad de expandir las funcionalidades de la plataforma para que se perciba más útil.

Tabla 1. Respuestas a la encuesta de conformidad con la tecnología ¹

Rol	Propuesta	MA	A	N	AD	MD
Generador de datos	Que tus datos sirvan de aportación para estudios de universidades y empresas en el mejoramiento de los sistemas de salud	10	1	1	0	1
	El uso de datos personales o de salud implica la firma de cartas de consentimiento. ¿Qué tan de acuerdo estás con hacerlo de manera electrónica a través de una app en tu celular por sobre la forma tradicional (carta impresa o manuscrita)?	8	2	2	0	1
	Conocer quién y con qué fin se usaron los datos que aportaste	9	2	1	0	1
	Conocer los resultados de los estudios en que se usaron tus datos	7	5	0	0	1
	Recibir remuneración o reconocimiento por tu aporte en estos estudios	9	4	0	0	0
Curador de datos	Que exista una trazabilidad de los datos usados en la confección del "dataset" hasta el usuario que los originó, para confirmar que provienen de fuentes confiables.	10	1	0	1	1
	Recibir remuneración o reconocimiento por tu trabajo en la confección del "dataset".	6	2	3	1	1
	Que exista un registro público de personas o instituciones que han consumido tu "dataset" así como con que finalidad lo hicieron. Todo esto con el fin de controlar el uso que se le da.	11	1	0	0	1
	El uso de datos personales o de salud requiere la firma de cartas de consentimiento. Como alguien que maneja este tipo de datos en la confección del "dataset": ¿Qué tan de acuerdo estás con hacerlo de manera electrónica a través de una app en tu celular por sobre la forma tradicional (carta impresa o manuscrita)?	7	3	0	2	1
Consumidor de datos	Que exista trazabilidad de los datos consumidos hasta el paciente origen con el fin de verificar que vienen de fuentes confiables (Data Provenance).	10	2	0	0	1
	Solicitar pacientes para estudios según criterios definidos (edad, etnia, nacionalidad, etc).	11	2	0	0	0
	Que se firmen cartas de consentimiento electrónicas en la plataforma para asegurar la no refutabilidad de los datos subidos y los derechos de acceso sobre los mismos.	11	1	0	0	1
	Que las fuentes de datos se clasifiquen por su confiabilidad. Ej: Fuentes de datos confirmadas por una institución contra fuentes no verificadas en Internet	9	3	0	0	1
	Poder solicitar el entrenamiento de modelos de aprendizaje de máquina con datos verificados por una institución y la correspondiente certificación del modelo	9	3	1	0	0
Totales		127	32	8	4	11
Porcentaje aproximado del total de respuestas (181)		70	17.6	4.3	2.1	6

¹MA: Muy de acuerdo, A: Algo de acuerdo, N: Neutral, AD: Algo en desacuerdo, MD: Muy en desacuerdo

El uso de un servidor web centralizado resulta un paso de retroceso si se desea obtener el nivel de descentralización esperado. La tecnología web usada en la interfaz no permite la compilación de contratos inteligentes para su despliegue y no accede a todos los sensores disponibles en el dispositivo móvil con la precisión que alcanzan las aplicaciones nativas.

Tomando estas enseñanzas se aboga por un enfoque más descentralizado, donde la aplicación cliente se conecte directamente a la cadena de bloques y los datos sean usados en operaciones de aprendizaje automático por un servicio independiente pero capaz de comunicarse con la arquitectura.

4.4. La implementación FRED

Partiendo del nuevo enfoque obtenido en 4.3 se realizó una implementación de la arquitectura TRHEAD enfocada en el sensado de imágenes de rostros humanos con el fin de entrenar un modelo de aprendizaje de máquina para la detección de estados de ánimo. Se realizó una campaña de sensado usando esta nueva plataforma, esta vez contando con funcionalidades dentro de la misma que permitieron capturar las imágenes y calificarlas entre válidas o no válidas, permitiendo que los participantes no solo tomaran el rol de generadores de datos, sino también el de curadores. La estructura de la plataforma; compuesta por una aplicación móvil, un contrato inteligente y la herramienta externa EvaNotebook; se describe a continuación.

4.4.1. Interfaz móvil

Con el fin de poder acceder a los sensores del dispositivo móvil para obtener una mayor precisión en las lecturas se optó por la tecnología Flutter para la implementación de la interfaz de usuario. La misma permite además crear vistas que se adaptan a la resolución de la pantalla y cuenta con librerías que permiten la comunicación directa con la Blockchain. La aplicación permite acceder a la cámara del dispositivo para tomar capturas, subirlas a un servidor IPFS y enviar al contrato inteligente los metadatos de la imagen y su CID asignado. Además, permite mostrar y etiquetar las imágenes subidas como válidas o no válidas para ser usadas en tareas de aprendizaje automático.

La aplicación cuenta con una vista para mostrar las estadísticas de las imágenes subidas y las vali-

daciones de las mismas y su uso implica la firma de una carta de consentimiento digital contenida en la vista de bienvenida.

4.4.1.1. Captura de imágenes

La vista de captura de imágenes de rostro que se puede apreciar en la Figura 11, permite que el usuario tome una foto de su tez o la seleccione de su galería. La imagen se subirá directamente al IPFS y se registrará mediante una llamada al contrato inteligente los metadatos de la misma como su nombre, tamaño, formato, identificador del usuario, emoción que refleja y su CID.



Figura 11. Captura de imágenes

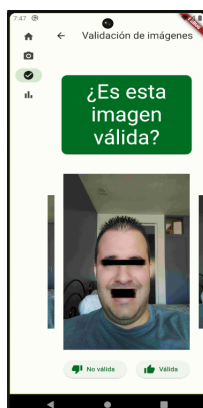


Figura 12. Validación de imágenes



Figura 13. Lista de imágenes

4.4.1.2. Vista de validación de imágenes

Las imágenes subidas a la plataforma usando la vista de captura se mostrarán en la vista de validación, dando la posibilidad a los usuarios de etiquetarlas como válidas o no válidas. El criterio a tener en cuenta es su utilidad en el entrenamiento de modelos de aprendizaje de máquina y para ello usan los botones que se pueden apreciar en la Figura 12. Ambas acciones agregan el identificador único del usuario que etiqueta a los metadatos de la imagen almacenados en el contrato inteligente, ya sea como defensor o detractor de su validez.

4.4.1.3. Listado de imágenes

En la vista de listado de imágenes expuesta en la Figura 13, se muestran las capturas almacenadas en la plataforma con la cantidad de usuarios que apoyan o están en contra de su validez. La finalidad de esto es apoyarse de una función de la propia interfaz para comparar la cantidad de votos a favor con aquellos que están en contra para etiquetar la imagen como válida o no válida. Similar al mecanismo de consenso conocido como “Proof of Stake”, aquellos usuarios que votaron incorrectamente se les penalizará retirándole créditos de la plataforma, los cuales se distribuirán entre aquellos que votaron correctamente. Esta última funcionalidad implicaba la creación de billeteras electrónicas y la gestión de tokens de la cadena de bloques donde se despliega el contrato inteligente. Al tratarse de una campaña voluntaria de sensado se optó por no implementar esta funcionalidad pues el motivo principal de estos tokens yace en la motivación, con la cual ya se contaba.

4.4.2. Contrato inteligente

El programa “Classifier” fue implementado con el fin de almacenar en la blockchain los metadatos de las imágenes clasificadas resultantes de la campaña de sensado. Se optó por usar la red Moonbeam para desplegar este contrato inteligente. Entre sus funcionalidades se encuentran las de agregar y eliminar nuevos registros de imágenes clasificadas, actualizar la listas de votos positivos y negativos, y borrar todos los registros, como medida en situaciones críticas. Un fragmento de esta implementación se puede observar en la Figura 14.

4.4.3. Entrenamiento de modelos predictivos

Las emociones en los individuos funcionan como indicadores fisiológicos y en la comunicación de información no verbal en general, existiendo evidencia de su universalidad (Tracy et al., 2015). La ira, el asco, el miedo, la alegría, la tristeza y la sorpresa reservan un estatus especial en todo el espectro de las emociones humanas porque han sido estudiadas ampliamente a lo largo de la historia y la literatura científica. Las expresiones faciales son una manifestación fisiológica de las emociones y gracias a estas es posible inferir la emoción que experimenta el individuo.

```

// SPDX-License-Identifier: MIT
pragma solidity 0.8.18;

contract ClassifiedData {
    struct ClassifiedImage {
        bytes32 cid;
        bytes32 emotion;
        bytes32 sourceUuid;
        bytes32 name;
        bytes32 size;
        bytes32[] supporters;
        bytes32[] opposers;
    }

    ClassifiedImage[] records;
    ClassifiedImage[] temp;

    function addRecord (bytes32 cid, bytes32 emotion,
        bytes32 sourceUuid, bytes32 name, bytes32 size) public {
        ClassifiedImage memory ci;
        ci.cid = cid;
        ci.emotion = emotion;
        ci.sourceUuid = sourceUuid;
        ci.name = name;
        ci.size = size;

        records.push(ci);
    }
}

```

Figura 14. Contrato inteligente para manejar metadatos de imágenes clasificadas

Por esto, proponemos realizar el proceso de “fine tuning” o afinado a una instancia de la arquitectura YOLO (You Only Look Once), una arquitectura unificada para la detección de objetos a través de imágenes en tiempo real. Para ello se usarán un conjunto de imágenes y las emociones asociadas a ellas, proporcionado por TRHEAD, para asegurar que son datos recolectados de manera consciente, con el fin de detectar expresiones faciales en las imágenes de rostros de pacientes.

4.4.3.1. Afinamiento

Las imágenes se obtuvieron de un directorio público en IPFS a través de su CID y otros metadatos necesarios como el estado de ánimo que representa cada una. Se realizó una exploración de datos usando el lenguaje Python como se indica en la Figura 15 y un pre-procesamiento usando la librería Pandas para redimensionar las imágenes a 640x640 píxeles y realizar su codificación entera, la cual permitió asignar un entero a cada clase de la siguiente manera:

- 0: Triste
- 1: Sorprendido
- 2: Feliz
- 3: Enojado

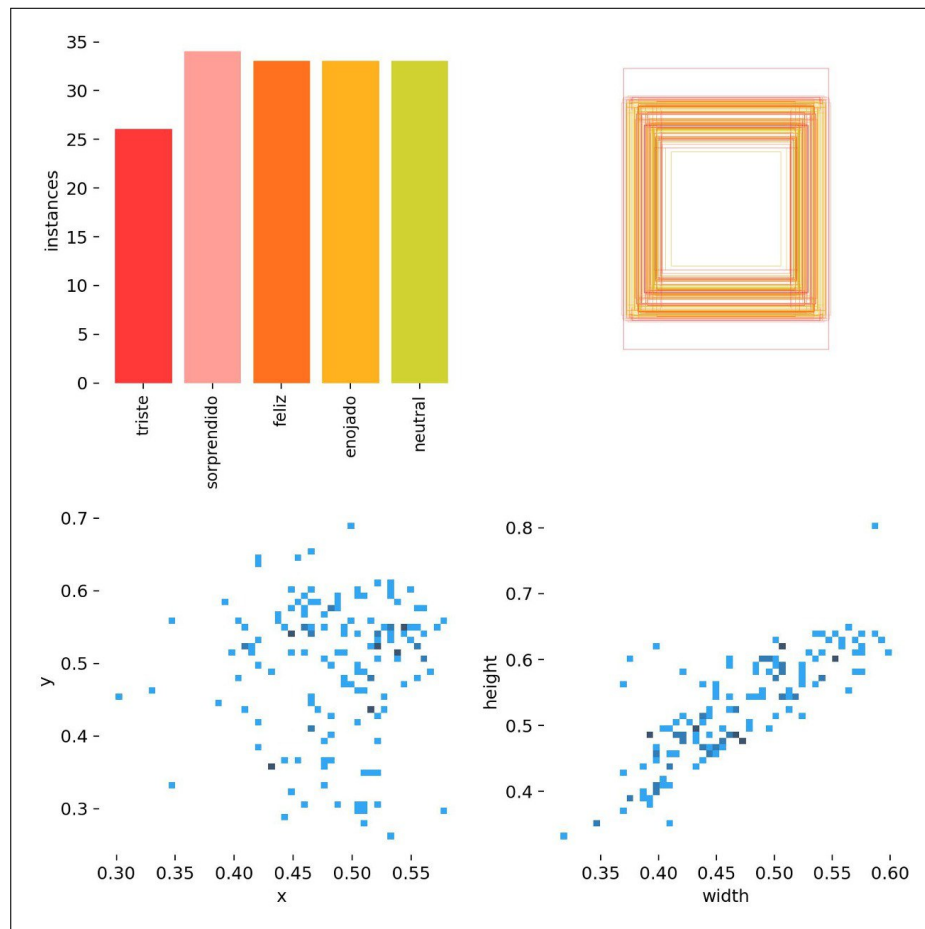


Figura 15. Distribución de las clases en el conjunto de imágenes

Finalmente se obtuvieron 159 instancias que cumplen con las especificaciones de YOLO en su versión 8, ajustándola con nuestro dataset con el fin de detectar emociones en el rostro de la persona dada una imagen. Para el entrenamiento se usaron 100 épocas, con 5 k-folds y se distribuyeron los datos de manera que se contara con una distribución de 80/20 por cierto para entrenamiento y validación, respectivamente.

Los resultados del modelo fueron prometedores, logrando un F1-Score, una métrica que combina precisión y recuperación, de 90.0% como se muestra en la Figura 16.

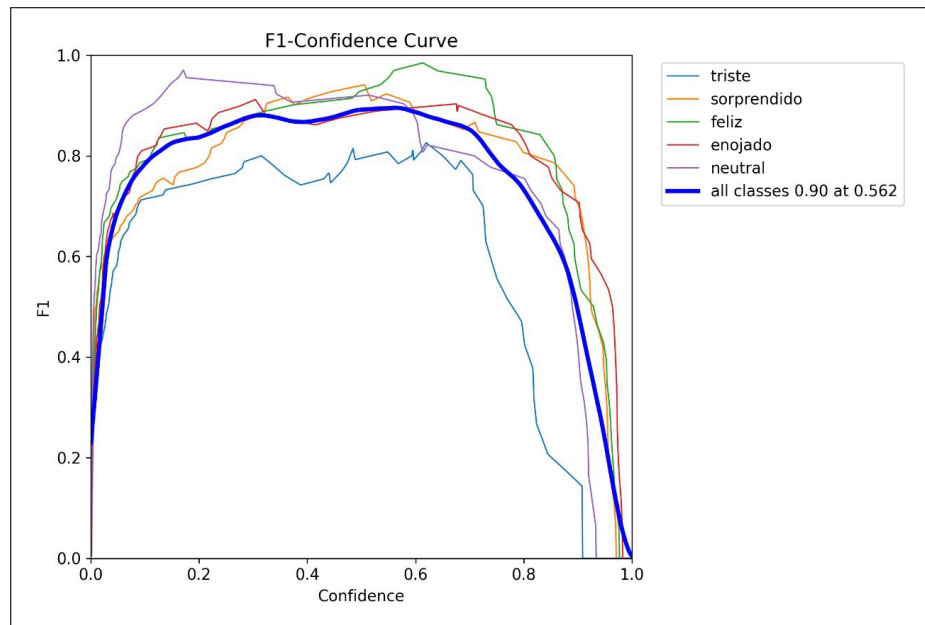


Figura 16. F1-Score del modelo entrenado

La curva de “Precisión-Recuperación” ², fue utilizada para evaluar la capacidad de clasificación del modelo en diferentes umbrales, mostró un área bajo la curva (AUC-PR) de 96.6 % expuesta en la Figura 17, lo que subraya la robustez del modelo en la clasificación de las categorías.

Estas métricas, en conjunto, demuestran la eficacia del modelo en la tarea para la que fue diseñado, aunque se observó una ligera tendencia al sobreajuste en las últimas épocas. Esto ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento, capturando tanto el patrón real de los datos como el ruido o la variabilidad aleatoria en los datos. Esto puede llevar a un rendimiento deficiente del modelo en datos nuevos o de prueba que no se han visto durante el entrenamiento.

En nuestro caso probablemente se debe a la cantidad limitada de datos de entrenamiento y podría corregirse con un mayor dataset o realizando validación cruzada, pero esto sería objeto de trabajo futuro. Podemos observar en la Figura 18 la matriz de confusión normalizada.

²La mayoría de los problemas de clasificación desequilibrada implican dos clases: un caso negativo con la mayoría de ejemplos y un caso positivo con una minoría de ejemplos. Se pueden crear gráficos a partir de las curvas y utilizarlos para comprender la compensación en el rendimiento para diferentes valores de umbral a la hora de interpretar predicciones probabilísticas. Cada gráfico también puede resumirse con una puntuación del área bajo la curva que puede utilizarse para comparar directamente los modelos de clasificación.

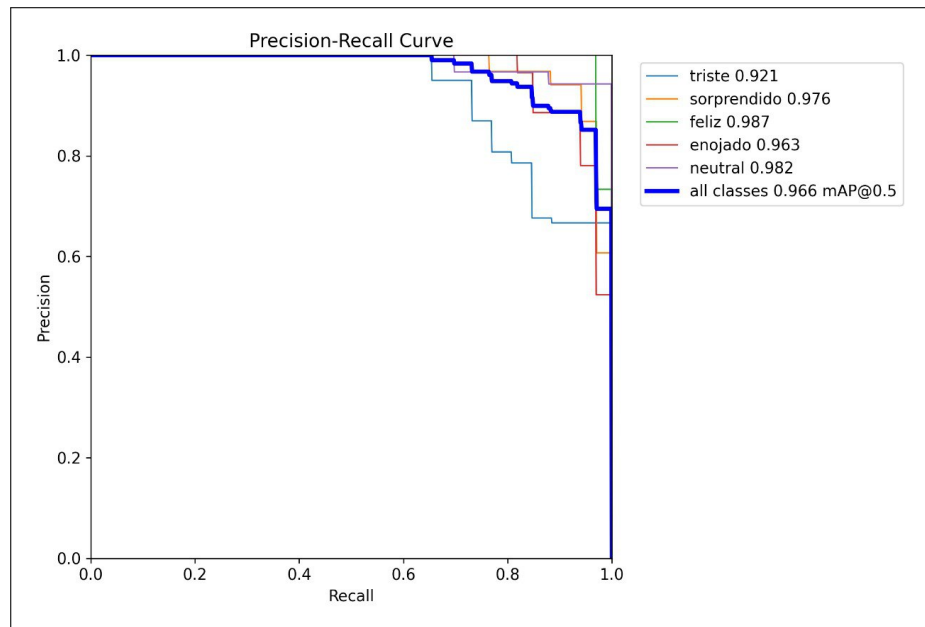


Figura 17. AUC-PR de la curva de “Precisión-Recuperación” que indica la capacidad de clasificación del modelo.

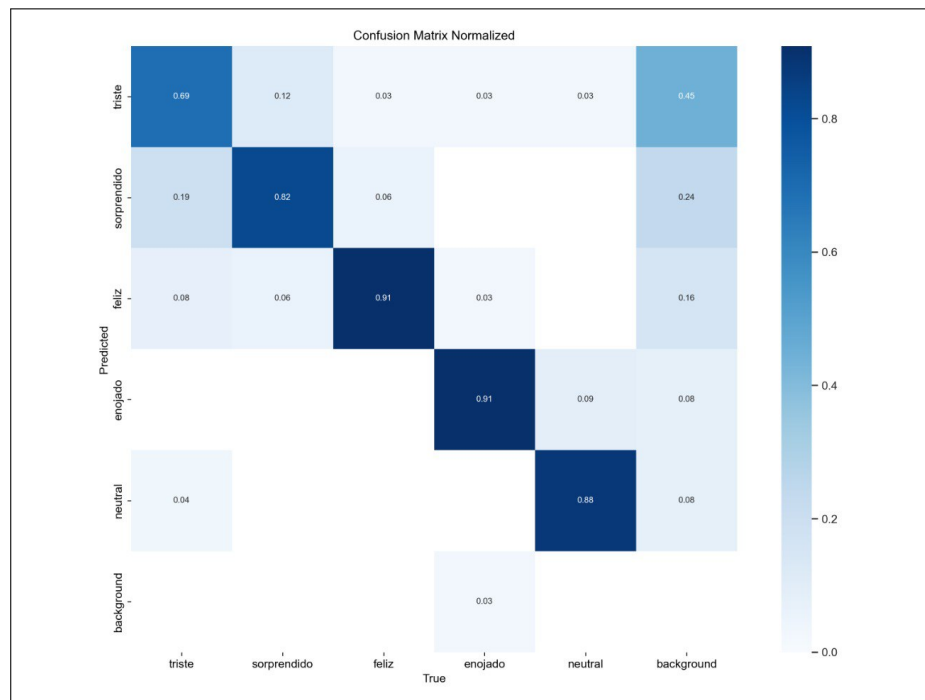


Figura 18. Matriz de confusión normalizada

Como se puede observar en la matriz de confusión, las clases “feliz y enojado” presentan la mayor sensibilidad y especificidad por lo que su clasificación será acertada en la mayoría de los casos, al contrario de “triste” que tiende a mezclarse con otras clases, incluso con el ruido que provoca el fondo de la imagen. La solución a esta clasificación incorrecta podría ser un preprocesamiento de las imágenes para eliminar

el ruido de fondo y el incremento de la presencia de la clase “triste” en los datos de entrenamiento.

El modelo resultante es capaz de identificar emociones a partir de una imagen según las clases establecidas durante su entrenamiento como se muestra en la Figura 19.

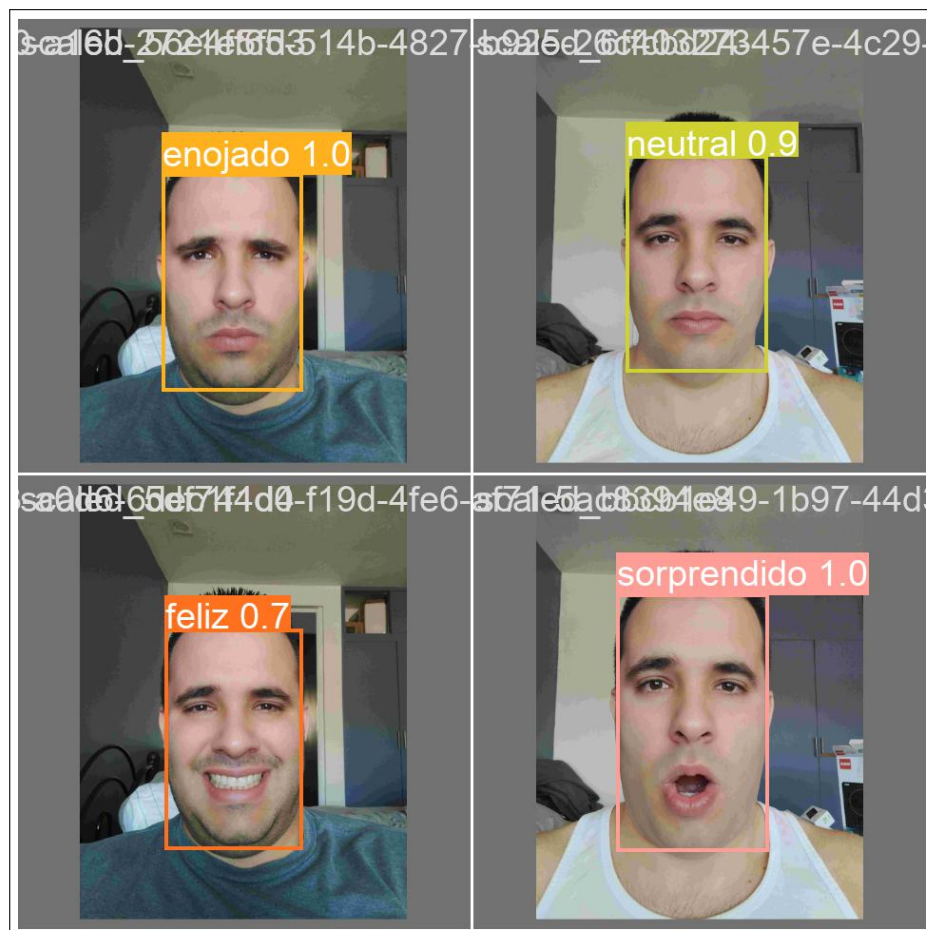


Figura 19. Ejemplo de detección de gestos faciales por el modelo

Con el fin de desplegar el modelo generado para su distribución y ejecución se adoptó EvaNotebook, desarrollada por Carlos Eduardo Sánchez Torres³, este componente será el encargado de recibir una imagen de un usuario y detectar la expresión facial que indica la emoción que imita el mismo.

³<https://orcid.org/0000-0001-5799-4067>

4.4.3.2. Despliegue en EvaNotebook

EvaNotebook es un cuaderno computacional diseñado específicamente para operar únicamente dentro de un entorno de navegador, sin necesidad de una arquitectura cliente-servidor, sino una arquitectura descentralizada “Peer to Peer” (P2P). Su diseño busca el desarrollo de prototipos, procesos experimentales y escenarios de sistemas con un énfasis principal en la programación multilenguaje para arquitecturas dirigidas por eventos. Esto se logra con la ayuda de una base de datos descentralizada e incorpora varios protocolos de aplicación, incluidos WebRTC, WebSockets y MQTT, lo que brinda un mayor nivel de abstracción.

En particular se usó EvaNotebook con THREAD para desplegar el modelo de la anterior sección de manera descentralizada. Un notebook primario descarga el modelo desde IPFS y lo ejecuta vía ONNX Runtime Web (un estándar de interoperabilidad de modelos), el notebook réplica llama al primario mediante WebRTC Data Communication y carga una imagen subida previamente a IPFS para su interpretación. Al primario se asigna un nombre de entidad desde el cual el resto de nodos puede ejecutar dicho modelo sin necesidad de instalarlo. De esta manera el secundario puede ejecutar el modelo para obtener una interpretación de la emoción que describe la imagen cargada y muestra la salida al usuario.

Habiendo expuesto cómo los datos recolectados por un sistema basado en TRHEAD pueden usarse en el entrenamiento de modelos de aprendizaje de máquina, el autor considera que se reúnen las condiciones para presentar al lector la discusión del trabajo realizado hasta el momento, a lo cual está dedicado el siguiente Capítulo 5.

Capítulo 5. Discusión

Habiendo expuesto los beneficios de una arquitectura basada en la Blockchain y las ventajas que esta ofrece sobre el enfoque tradicional, podría dar la impresión de ser la herramienta ideal en la recopilación y uso consciente y legal de datos de sensado móvil. Sin embargo, está claro que este tipo de tecnologías dependen en gran medida de la participación masiva de los usuarios. La estabilidad y seguridad de la propia cadena dependen del número de nodos que se integran en ella y el futuro de esta, a su vez, depende de las intenciones de la mayoría de los usuarios.

5.1. Desafíos técnicos

A pesar del enfoque descentralizado de las blockchain, los mecanismos de consenso que implementan pueden ser manipulados para centralizar el control de la red en un grupo con intenciones unilaterales. Por ejemplo, “Proof of Stake” tiende a dar mayor peso a los validadores con mayor cantidad de participación en la cadena, lo que se traduce en que a mayor cantidad de tokens puestos en juego, mayor probabilidad de ser elegido como validador. El principal problema en este caso lo constituyen los mercados y pasarelas de pago de criptomonedas, que mediante la realización de sus actividades llegan a tener un “stake” tan alto que aumenta las posibilidades de ser seleccionados como validadores.

Si varias de estas instituciones formasen una coalición podrían centralizar sobre sí las validaciones de los nuevos bloques de una cadena, tomando efectivamente el control sobre la misma. Ciertamente esto solo se cumple para aquellas cadenas que implementan “Proof of Stake”, aquellas que implementan “Proof of Work” no tienen esta vulnerabilidad pero tienen serios problemas de rendimiento en comparación con aquellas que usan mecanismos de consenso más recientes. Es necesario resaltar las consecuencias ambientales que tiene el minado de criptomonedas por el alto consumo energético de los equipos y el impacto que su compra masiva tiene en el mercado (Badea & Mungiu-Pupazan, 2021).

Otro problema ligado a estas arquitecturas reside precisamente en una de sus ventajas, el carácter democrático de la red, que transforma la toma de decisiones sobre el funcionamiento de la misma en un proceso participativo. En casos muy especiales en que los nodos participantes no llegan a un acuerdo que satisfaga a la mayoría ocurre un “fork”, proceso mediante el cual se divide la red en dos redes independientes que llegado cierto punto pueden perder la interoperabilidad entre sí. Esto podría usarse a modo de ataque adversario para aislar un grupo de nodos de la red principal (Liao & Katz, 2017). En

un caso hipotético de varias instituciones de salud participando de un sistema basado en la Blockchain “X”, si ocurriese un “fork” de la misma en “X1” y “X2”, se reduciría la cantidad de nodos en cada red, afectando su robustez y seguridad. Incluso, si la cadena “X1” recibe cambios en su implementación que la separan demasiado de la original, se perderá la posibilidad de comunicación con “X2” haciendo que los nuevos registros de una no puedan ser replicados en la otra. Hay Blockchains como Polkadot ¹ que promueven una estructura y mecanismos de consenso más centrados en la retrocompatibilidad entre subcadenas o parachains que prometen ser la solución más inmediata a este reto. Al desprenderse de un tronco común conocido como Substrate ², la compatibilidad entre sus parachains es inherente a su propia construcción, lo que les permite contar con interoperabilidad entre partes luego de un fork e incluso entre otras parachains, al punto de que todas manejan el mismo tipo de billetera electrónica.

Otra dificultad que enfrenta la introducción de arquitecturas de sistemas basados en la “cadena de bloques” es la complejidad técnica para el usuario final del manejo de activos digitales, billeteras electrónicas y transacciones con monedas que desconocen y cuyo precio sufre cambios bruscos. También para los desarrolladores resulta complicado introducirse al desarrollo de contratos inteligentes por la pronunciada curva de aprendizaje de lenguajes como Solidity y las consecuencias de una implementación incorrecta de los mismos teniendo en cuenta que manejan activos y no están pensados para almacenar información.

En respuesta a esto han surgido compañías como la Fundación Gear³, enfocada en ofrecer capacidades de Web3 al tiempo que garantiza una experiencia de usuario similar a Web2. Introduce una mayor velocidad, seguridad y automatización, aportando avances significativos al desarrollo de contratos inteligentes como plantillas personalizables de contratos inteligentes y compatibilidad con WASM⁴, para el despliegue de sistemas usando su cadena Vara Network⁵, una parachain de Polkadot. Esta orientación a propiciar la aceptación de la tecnología permite que más desarrolladores puedan optar por aprovechar las propiedades de la Blockchain en sus aplicaciones de servicios de salud.

La inmutabilidad de los contratos inteligentes es la razón principal de su uso como mediador para el consenso entre dos partes sin necesidad de que existan confianza mutua. El hecho de no poder cambiar los estatutos del contrato y de que las operaciones solo se realizan una vez se cumplan las condiciones establecidas en el código, ofrece la confianza que generalmente aporta un tercero en este tipo de

¹<https://www.polkadot.network/>

²<https://substrate.io/>

³<https://gear-tech.io/>

⁴<https://webassembly.org/>

⁵<https://vara-network.io/>

transacciones. Sin embargo, en esto reside su principal problema, ya que en caso de existir un error en el código o se deseen modificar los estatutos por mutuo acuerdo de los pares, no existe una vía trivial para realizar una corrección (Khan et al., 2021). Existen enfoques para solucionar estas situaciones como realizar un proceso de “proxy” a un contrato, mediante el cual se redirige la versión que se desea cambiar del contrato a una nueva versión previamente desarrollada, cabe aclarar que este “proxy” debió haber sido definido en el contrato original antes de su despliegue.

5.2. Desafíos Éticos

La legislación de un proceso pretende regular los métodos utilizados para el desarrollo del mismo y el manejo de datos no es una excepción. A menudo, las medidas que se adoptan contra los infractores de una legislación giran en torno a multas o procedimientos judiciales contra los responsables de su violación. Sin embargo, cuando las pérdidas económicas que esto supone son inferiores al beneficio total que reporta la propia infracción, lo que debería ser una reprimenda, se vuelve un negocio rentable. Por otro lado, el exceso de presión legal entorpece la labor de las empresas que están dispuestas a respetar los derechos de los usuarios y sólo desean ofrecer un servicio de calidad. Una respuesta a esta encrucijada podría encontrarse en el uso de la arquitectura de referencia TRHEAD para abordar algunos de estos desafíos éticos.

5.2.1. Uso justo de los datos

La ética de aquellos a quienes se les confirió la protección de nuestros datos desempeña un papel fundamental en el tratamiento justo de los mismos. El acuerdo de términos de servicio no puede dar a una empresa rienda suelta sobre los datos de un usuario indefinidamente y no se puede recopilar ningún dato hasta que se haya dado el consentimiento para ese fin, pudiendo este retractarse en cualquier momento. Contratos inteligentes como los que implementan los sistemas basados en TRHEAD podrían automatizar esta tarea al almacenar las cartas de consentimiento de los usuarios en NFTs.

Las funciones propias de estos contratos permiten a su poseedor el control de la vigencia, cantidad de copias disponibles, propiedad y estado en la red de sus activos, además de convertirlos en un objeto

único y asignarles un valor dentro de la cadena. Los NFT son transferibles por lo que representan una versión digital de la carta de consentimiento que puede ser comercializada como garantía de la procedencia de los datos. Al ser parte de la Blockchain todas las transacciones que se realizan sobre esta quedan registradas en los bloques, aportando trazabilidad de los datos y permitiendo al paciente conocer la finalidad de sus registros médicos.

5.2.2. Protección de la privacidad

Otro enfoque del problema consiste en garantizar la protección de la identidad del usuario en lugar de la seguridad de sus datos, de modo que aunque los datos se viesan comprometidos no habría forma directa de vincularlos al paciente. De lograr este acometido contaríamos con datos anónimos, listos para su consumo en el desarrollo de sistemas para el cuidado de la salud, sin embargo los consumidores necesitan verificar la procedencia de estos.

La procedencia implica saber que fueron recopilados legal y conscientemente, que fueron correctamente procesados y etiquetados, y que no han sido alterados o generados de forma sintética(Shumailov et al., 2023). En un sistema que cumpla con los requisitos de TRHEAD, el proceso de recogida, tratamiento y consumo de datos quedaría registrado en la Blockchain, aprovechando su inmutabilidad y mecanismo de consenso para garantizar que los registros son válidos. La identidad del usuario sería sustituida por una dirección de monedero electrónico que le identificará en la red, así como las transacciones que realice sobre su propiedad virtual (datos del paciente), añadiendo un componente de no refutabilidad a sus operaciones como la firma de una carta de consentimiento.

5.2.3. Sesgo

El sesgo se produce cuando los sujetos son “elegidos a mano” durante la recogida de datos. Algunas personas cometen el error de pensar que un poco de sesgo es inofensivo, pero podría afectar al resultado final de un determinado modelo de aprendizaje automático introduciendo desequilibrios en los datos y errores de medición causados por un muestreo no representativo de la población (Mehrabi et al., 2021).

Una implementación de TRHEAD con el apoyo apropiado podría tener alcance internacional, contando con usuarios de distintas etnias, nacionalidades, edades, sexo, etc. El carácter participativo de la red atraería a nuevos pacientes al sensado activo de su estado de salud, contribuyendo no solo a la cantidad sino a la variedad de los datos. La estructura distribuida facilita el acceso a la plataforma siempre que existan nodos de la misma en la zona desde donde se desean consumir sus servicios, a la vez que asegura su disponibilidad. De esta forma los modelos de aprendizaje de máquina resultantes serían entrenados con datos representativos, obteniendo mayor robustez y precisión.

5.2.4. Cultura de la justeza

El uso justo de los datos no sólo debe estar presente en el tratamiento de los mismos, sino que debe considerarse como un proceso económico-social y regularse como tal. En este proceso intervienen los roles principales de la arquitectura TRHEAD 3.1 como usuarios. El autor considera justo que los derechos y deberes de estos usuarios se distribuyan como se muestra en la Tabla 2. El hecho de que la responsabilidad del uso correcto de los datos recaiga en los roles antes mencionados no implica que no deban responder a las autoridades pertinentes.

Se puede contar con una red manejada democráticamente por sus participantes y a la vez regulada por una institución externa, evitando así operaciones dentro de la misma que atenten contra su estabilidad. Aunque pueda parecer que esto concierne solo a los usuarios de la misma, quienes tienen el poder para controlarla, no es correcto ver una Blockchain como un ambiente aislado, sino como parte del sistema económico, puesto que eventos fuera de la red pueden afectar el precio de los activos que en esta se manejan. De igual manera la inflación que puede darse en el precio de los tokens de una Blockchain puede afectar el precio de artículos fuera de la misma (Thomson, 2020).

Partiendo de los puntos discutidos en esta sección, estamos en condiciones de arribar a conclusiones sobre el uso de la arquitectura de referencia TRHEAD en la recopilación y uso justo de datos de sensado móvil y vestible.

Tabla 2. Derechos y deberes de los usuarios según su rol

Roles	Derechos	Deberes
Generadores y tenedores de datos	Ser recompensados por su contribución en forma de moneda, servicio o reconocimiento.	Aceptar la no refutabilidad de las operaciones que realizan sobre sus datos.
	Respeto a su privacidad.	Respetar las normativas de los estudios de salud en los que participan.
	Controlar el acceso a sus datos y su "presencia en la red".	
Curadores de datos	Ser recompensados por su contribución en forma de moneda, servicio o reconocimiento.	Aceptar la no refutabilidad de las operaciones que realizan sobre los datos de los usuarios y en su propio trabajo.
	Tener su trabajo reconocido como un activo independiente de los datos utilizados para producirlos.	Respetar la privacidad de los usuarios incluso cuando los datos son proporcionados por los tenedores.
	Rastrear los datos hasta fuentes confiables para verificar su procedencia.	Requerir el uso de mecanismos legales como cartas de consentimiento bajo la aprobación de un comité de ética.
Consumidores de datos	Poder verificar la procedencia de los datos que están consumiendo.	Asegurarse de que los datos que reciben fueron recopilados conscientemente y bajo la autorización de los pacientes.
	Tener sus derechos sobre los datos consumidos listados y reconocidos a través de algún mecanismo.	Usar los datos bajo los términos acordados con el usuario a través de cartas de consentimiento o similares.
	Recibir los datos bajo los términos acordados durante la transacción, en el caso de estudios médicos patrocinados y similares.	

Capítulo 6. Conclusiones

Vivimos en la era de la información y las comunicaciones, donde los avances tecnológicos son omnipresentes en nuestro día a día e influyen en nuestra forma de percibir el mundo que nos rodea. La materia prima de la información son los datos, aquellos que producimos como resultado de nuestra interacción con el medio que nos rodea. Estos datos dicen quienes somos y se podrían utilizar para mejorar los servicios que nos ofrecen, como la atención médica, siempre y cuando se les de el tratamiento adecuado.

Habiendo llegado a término de la labor investigativa correspondiente a este trabajo de tesis, hemos analizado tendencias en el uso de datos de sensado móvil y vestible para su aplicación en los servicios sanitarios. Se propuso la arquitectura de referencia TRHEAD para la implementación de sistemas que realicen esta tarea asegurando la procedencia de los datos sin comprometer la privacidad del paciente. Se expuso como la Blockchain es una plataforma que, a través de los contratos inteligentes, asegura la trazabilidad y no refutabilidad de las transacciones que en ella se realizan y se exploraron sus aplicaciones en el campo de la salud. Se demostró que es posible consumir los datos legal/conscientemente recopilados por sistemas basados en la arquitectura TRHEAD para el entrenamiento de modelos de aprendizaje de máquina. Hemos descrito cómo los NFTs pueden convertirse en una versión digital de una carta de consentimiento, permitiendo a los pacientes controlar y rastrear el uso de sus datos. Entendemos que el camino a la recolección ética de datos no debería ser el de proteger a los usuarios de cada nueva amenaza que aparece, sino darle las herramientas para defender sus derechos y tomar control de su identidad en la red.

6.1. Contribuciones

La realización de este trabajo tuvo como resultado el diseño de la arquitectura de referencia TRHEAD para sistemas de recolección y uso de datos de sensado móvil y vestible. De esta arquitectura se realizó una implementación inicial, cuya fase de prueba no solo demostró su viabilidad sino la existencia de un interés legítimo en la conservación de su privacidad por parte de los usuarios y conciencia del rol que deben jugar los consumidores de datos en el respeto a la misma. Una segunda implementación demostró cómo los sistemas basados en TRHEAD se pueden beneficiar de una arquitectura más orientada a la descentralización para no adolecer de los problemas que afectan a los servicios hospedados en servidores unificados.

En adición a esto se consiguió conectar la plataforma a una herramienta externa para la realización

de tareas de MLOps llamada EvaNotebook, donde se ejecuta un modelo pre-entrenado para la detección de emociones en una imagen estática. Como resultado de la labor investigativa llevada cabo se redactó el artículo “Smart Contracts for the Fair Treatment of Mobile Sensing Data”; de los autores Dr. Jesús Favela Vara, Carlos Eduardo Sánchez Torres y un servidor; aceptado para su publicación en el foro Workshop Fair’23 de la Conferencia de Cómputo Ubicuo y el Simposio Internacional de Cómputo Vestible UBICOMP-ISWC en su edición del presente año 2023. De los mismos autores se encuentra en proceso de evaluación el artículo “Traceable Health Data for Consciously Trained ML models”, presentado para en la 15ta Conferencia Internacional de Cómputo Ubicuo y Ambientes Inteligentes (UCAml’23).

6.2. Trabajo a futuro

Entendemos que aún queda mucho trabajo por hacer en aras de ofrecer un servicio óptimo que atraiga el volumen de usuarios esperado para una amplia difusión de la plataforma. Comprendemos la inquietud de muchos pacientes sobre la protección de sus datos, por lo que proponemos agregar un enfoque de encriptación asimétrica de los mismos. Funcionaría escondiendo una clave secreta dentro de los archivos a modo de marca de agua, la cual se encripta mediante el Estándar de Encriptación Avanzada (AES), estos archivos marcados se firman con la clave pública del usuario y también son encriptados justo antes de subirlos al almacenamiento público.

Las cartas de consentimiento digitales de estos datos contendrán la clave para descifrarlos. Este enfoque se recomendaría para información que no sea de dominio público o aquella directamente destinada a la comercialización. Para mejorar la integración de la plataforma con servicios de terceros se plantea la implementación de un contrato inteligente encargado de almacenar los registros de datos disponibles para MLOps como un archivo JSON y generarlo para consumo a través de una llamada desde la aplicación cliente.

Literatura citada

- Al-Janabi, S., Al-Shourbaji, I., Shojafar, M., & Shamshirband, S. (2017). Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications. *Egyptian Informatics Journal*, 18, 113–122. <https://doi.org/10.1016/J.EIJ.2016.11.001>.
- Alharby, M. & van Moorsel, A. (2017). Blockchain-based smart contracts: A systematic mapping study. *Fourth International Conference on Computer Science and Information Technology*, 125–140. <https://doi.org/10.5121/csit.2017.71011>.
- Angelov, S., Grefen, P., & Greefhorst, D. (2012). A framework for analysis and design of software reference architectures. *Information and Software Technology*, 54, 417–431. <https://doi.org/10.1016/J.INFSOF.2011.11.009>.
- Aste, T., Tasca, P., & Matteo, T. D. (2017). Blockchain technologies: The foreseeable impact on society and industry. *Computer*, 50, 18–28. <https://doi.org/10.1109/MC.2017.3571064>.
- Badea, L. & Mungiu-Pupazan, M. C. (2021). The economic and environmental impact of bitcoin. *IEEE Access*, 9, 48091–48104. <https://doi.org/10.1109/ACCESS.2021.3068636>.
- Bardram, J. E. (2008). Pervasive healthcare as a scientific discipline. *Methods of information in medicine*, 47, 178–185. <https://doi.org/10.3414/ME9107>.
- Biørn-Hansen, A., Majchrzak, T. A., & Grønli, T.-M. (2017). Progressive web apps: The possible web-native unifier for mobile development. In *Proceedings of the 13th International Conference on Web Information Systems and Technologies - Volume 1: WEBIST*, 344–351. INSTICC, SciTePress. <https://doi.org/10.5220/0006353703440351>.
- Dunning, A., Smaele, M. D., & Böhmer, J. (1970). Are the fair data principles fair? *International Journal of Digital Curation*, 12, 177–195. <https://doi.org/10.2218/IJDC.V12I2.567>.
- Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research*, 3, 1–5. <https://doi.org/10.2196/jmir.3.2.e20>.
- Favela, J., Castro, L. A., & Michan, L. (2016). Towards a federated repository of mobile sensing datasets for pervasive healthcare. *International Conference on Pervasive Computing Technologies for Healthcare*. <https://doi.org/10.5555/3021319.3021373>.
- Feng, X., Shen, J., & Fan, Y. (2009). Rest: An alternative to rpc for web services architecture. In *2009 First International Conference on Future Information Networks*, 7–10. <https://doi.org/10.1109/ICFIN.2009.5339611>.
- Florentino, L., Rissling, M., Liu, L., & Ancoli-Israel, S. (2011). The symptom cluster of sleep, fatigue and depressive symptoms in breast cancer patients: Severity of the problem and treatment options. *Drug Discovery Today: Disease Models*, 8, 167–173. <https://doi.org/10.1016/j.ddmod.2011.05.001>.
- Grouios, G., Ziagkas, E., Loukovitis, A., Chatzinikolaou, K., & Koidou, E. (2023). Accelerometers in our pocket: Does smartphone accelerometer technology provide accurate data? *Sensors (Basel, Switzerland)*, 23. <https://doi.org/10.3390/S23010192>.
- Haleem, A., Javaid, M., Singh, R. P., Suman, R., & Rab, S. (2021). Blockchain technology applications in healthcare: An overview. *International Journal of Intelligent Networks*, 2, 130–139. <https://doi.org/10.1016/J.IJIN.2021.09.005>.
- Hassanein, A. A., El-Tazi, N., & Mohy, N. N. (2022). Blockchain, smart contracts, and decentralized applications: An introduction. *Blockchain technologies*, 97–114. https://doi.org/10.1007/978-981-16-3412-3_6.

- Howell, S. T., Niessner, M., & Yermack, D. (2020). Initial coin offerings: Financing growth with cryptocurrency token sales. *The Review of Financial Studies*, 33, 3925–3974. <https://doi.org/10.1093/RFS/HHZ131>.
- Hölbl, M., Kompara, M., Kamišalić, A., & Zlatolas, L. N. (2018). A systematic review of the use of blockchain in healthcare. *Symmetry* 2018, Vol. 10, Page 470, 10, 470. <https://doi.org/10.3390/SYM10100470>.
- Kalnis, P., Ghinita, G., Mouratidis, K., & Papadias, D. (2007). Preventing location-based identity inference in anonymous spatial queries. *IEEE Transactions on Knowledge and Data Engineering*, 19, 1719–1733. <https://doi.org/10.1109/TKDE.2007.190662>.
- Khan, S. N., Loukil, F., Ghedira-Guegan, C., Benkhelifa, E., & Bani-Hani, A. (2021). Blockchain smart contracts: Applications, challenges, and future trends. *Peer-to-Peer Networking and Applications*, 14, 2901–2925. <https://doi.org/10.1007/S12083-021-01127-0/FIGURES/4>.
- Kotz, D., Gunter, C. A., Kumar, S., & Weiner, J. P. (2016). Privacy and security in mobile health: A research agenda. *Computer*, 49, 22–30. <https://doi.org/10.1109/MC.2016.185>.
- Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12, 74–82. <https://doi.org/10.1145/1964897.1964918>.
- Leonelli, S., Lovell, R., Wheeler, B. W., Fleming, L., & Williams, H. (2021). From fair data to fair data use: Methodological data fairness in health-related social media research. <https://doi.org/10.1177/20539517211010310>, 8. <https://doi.org/10.1177/20539517211010310>.
- Liao, K. & Katz, J. (2017). Incentivizing blockchain forks via whale transactions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10323 LNCS, 264–279. https://doi.org/10.1007/978-3-319-70278-0_17/TABLES/3.
- Mann, S. P., Savulescu, J., Ravaud, P., & Benchoufi, M. (2021). Blockchain, consent and present for medical research. *Journal of Medical Ethics*, 47, 244–250. <https://doi.org/10.1136/MEDETHICS-2019-105963>.
- Marangunić, N. & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, 14, 81–95. <https://doi.org/10.1007/S10209-014-0348-1/TABLES/3>.
- McGhin, T., Choo, K. K. R., Liu, C. Z., & He, D. (2019). Blockchain in healthcare applications: Research challenges and opportunities. *Journal of Network and Computer Applications*, 135, 62–75. <https://doi.org/10.1016/J.JNCA.2019.02.027>.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54. <https://doi.org/10.1145/3457607>.
- Omar, I. A., Jayaraman, R., Salah, K., Simsekler, M. C. E., Yaqoob, I., & Ellahham, S. (2020). Ensuring protocol compliance and data transparency in clinical trials using blockchain smart contracts. *BMC medical research methodology*, 20. <https://doi.org/10.1186/S12874-020-01109-5>.
- Pearson, S., May, D., Leontidis, G., Swainson, M., Brewer, S., Bidaut, L., Frey, J. G., Parr, G., Maull, R., & Zisman, A. (2019). Are distributed ledger technologies the panacea for food traceability? *Global Food Security*, 20, 145–149. <https://doi.org/10.1016/j.gfs.2019.02.002>.

- Ranga, V. & Soni, A. (2019). Api features individualizing of web services: Rest and soap actuator design and testing for the deployment of fire extinguisher ball using iot based drone view project partition recovery of wireless sensor networks with the integration of unmanned aerial vehicles (uavs) view project. *Article in International Journal of Innovative Technology and Exploring Engineering*, 2278–3075. <https://doi.org/10.35940/ijitee.I1107.0789S19>.
- Rauchs, M., Glidden, A., Gordon, B., Pieters, G. C., Recanatini, M., Rostand, F., Vagneur, K., & Zhang, B. Z. (2018). Distributed ledger technology systems: A conceptual framework. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.3230013>.
- Ray, P. P., Dash, D., & De, D. (2019). A systematic review and implementation of iot-based pervasive sensor-enabled tracking system for dementia patients. *Journal of Medical Systems*, 43, 1–21. <https://doi.org/10.1007/S10916-019-1417-Z/TABLES/8>.
- Sahi, M. A., Abbas, H., Saleem, K., Yang, X., Derhab, A., Orgun, M. A., Iqbal, W., Rashid, I., & Yaseen, A. (2017). Privacy preservation in e-healthcare environments: State of the art and future directions. *IEEE Access*, 6, 464–478. <https://doi.org/10.1109/ACCESS.2017.2767561>.
- Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The curse of recursion: Training on generated data makes models forget. *arXiv Labs*. <https://doi.org/https://doi.org/10.48550/arXiv.2305.17493>.
- Sridhar, Suman, A., Adari, K., Alla, S., & Adari, S. K. (2021). What is mlops? *Beginning MLOps with MLFlow*, 79–124. https://doi.org/10.1007/978-1-4842-6549-9_3.
- Taherdoost, H. (2022). Non-fungible tokens (nft): A systematic review. *Information 2023, Vol. 14, Page 26*, 14, 26. <https://doi.org/10.3390/INF014010026>.
- Thomson, J. C. (2020). *Tragedy of the Energy Commons: How Government Regulation Can Help Mitigate the Environmental and Public Health Consequences of Cryptocurrency Mining*. Seattle Journal of Technology, Environmental and Innovation Law (SJTEIL).
- Torre, I., Koceva, F., Sanchez, O. R., & Adorni, G. (2016). A framework for personal data protection in the iot. 384–391. IEEE. <https://doi.org/10.1109/ICITST.2016.7856735>.
- Tracy, J. L., Randles, D., & Steckler, C. M. (2015). The nonverbal communication of emotions. *Current Opinion in Behavioral Sciences*, 3, 25–30. Social behavior, <https://doi.org/https://doi.org/10.1016/j.cobeha.2015.01.001>.
- Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, 12, 113–127. <https://doi.org/10.1007/S11036-007-0017-1/METRICS>.
- Vijayalakshmi, K., Bushra, S. N., Subramanian, N., & Ponnuramu, V. (2022). Blockchain based medical record storage and retrieval using nft tracking system. *2022 6th International Conference on Trends in Electronics and Informatics, ICOEI 2022 - Proceedings*, 795–802. <https://doi.org/10.1109/ICOEI53556.2022.9776833>.
- Weiss, K., Khoshgoftaar, T. M., & Wang, D. D. (2016). A survey of transfer learning. *Journal of Big Data*, 3, 1–40. <https://doi.org/10.1186/S40537-016-0043-6/TABLES/6>.
- Zghaibeh, M., Farooq, U., Hasan, N. U., & Baig, I. (2020). Shealth: A blockchain-based health system with smart contracts capabilities. *IEEE Access*, 8, 70030–70043. <https://doi.org/10.1109/ACCESS.2020.2986789>.
- Zhang, J., Xue, N., & Huang, X. (2016). A secure system for pervasive social network-based healthcare. *IEEE Access*, 4, 9239–9250. <https://doi.org/10.1109/ACCESS.2016.2645904>.