Formato A1

Premio de Innovación y Buenas Prácticas en Protección de Datos Personales 2024

Trazabilidad en procesos de donación de datos médicos

Autores

Nombre de la persona con la que el Comité técnico establecerá contacto durante el proceso de recepción de trabajos y de evaluación de forma.

*Dra. María de los Ángeles Cosío León**

Dra. Anabel Martínez Vargas

INDICE

Desarrollo del Resumen Ejecutivo

Certificar, es proveer un documento que acredite que una persona, producto, sistema o servicio cumple con ciertos estándares o requisitos, y para poder acreditarlo se requieren de elementos que demuestren su cumplimiento sin duda legal o técnica.

La trazabilidad es el proceso mediante el cual se puede rastrear el historial, la ubicación o la trayectoria de un producto, servicio o proceso a lo largo de su ciclo de vida.

Lo que se propone es un mecanismo que sigue y registra cada una de las acciones de un donante de datos clínicos que lo llevaran a ceder con éxito la propiedad de sus datos clínicos o bien abortar dicho proceso. esto aumenta su importancia cuando la información proviene de grupos vulnerables, como el caso de menores de edad. Este mecanismo le dará la capacidad de revisar el tratamiento de sus datos personales y de su información sensible en el ecosistema de datos, destino de su donación, esto durante el proceso o después de concluida la interacción. Para ello la primera acción del donante es generar un registro como donante en el sistema, lo que permite asociar a un donante con un identificador de proceso de donación e iniciar el ingreso de los datos clínicos.

Al ingresar los datos se inicia la traza que comprende las diversas etapas del proceso de donación. a) identificación de la fuente de datos (humanos, en diferentes edades y condición física; o sintéticos), b) adquisición de datos clínicos, c) anonimización de datos clínicos para el cumplimiento de la normatividad legal vigente en materia de datos clínicos, y d) su almacenamiento seguro, si el proceso se concluye de estar manera o la eliminación de la información asociada al proceso de donación en cuestión. Que finalmente se registraran cada una de ellas en la blockchain y le envía un código QR al donador, código que le permite navegar en cada uno de los pasos ejecutados por el sistema y los códigos de programación (contratos inteligentes) que se usaron para ejecutar las acciones, ingreso de datos-anonimización.

Después de este punto, ya no le es posible al donante ver en el sistema los datos clínicos que dono, o que opto por no hacerlo. La razón es porque en caso de que el proceso de donación se concluyera con éxito, los datos ya no tienen información que permita su relación con el donante, condición que se indica en el formato de donaciones **Anexo A**; o en caso de que la donación no se realizara, todos los archivos son eliminados del sistema.

Objetivo de la buena práctica o del elemento innovador

OBJETIVO: Generar l*a traza de un proceso de donación de datos clínicos por personas físicas,* ***mediante*** *técnicas de blockchain, este proceso incluye las siguientes etapas: a) identificación de la fuente de datos (humanos, en diferentes edades y condición física; o sintéticos), b) adquisición de datos clínicos, c) anonimización de datos clínicos para el cumplimiento de la normatividad legal vigente en materia de datos clínicos, y d) su almacenamiento seguro* ***que nos permitan*** *contar con certeza legal y técnica para su uso por entidades tanto físicas como morales en el entrenamiento de modelos de inteligencia artificial y personal de área médicas; con el objetivo de apoyar en la detección de enfermedades en la población mexicana.*

Desarrollo de la innovación o buena práctica propuesta

En un primer paso el donante debe registrase, para ello se propone usar un servicio IAM, que permitirá que al donante se les asignen roles y acceso a servicios relacionados con el proceso de donación. La intervención del donante es mínima y será agregado al grupo donadores. Una vez registrado, el usuario firma un Formulario de Consentimiento (Anexo A) en el cual se indica el proceder sobre sus datos personales y que posteriormente él podrá verificar al acceder al sistema que le permitirá hacer una comparativa de los datos que entrego y su estado una vez anonimizados, punto en el cual puede verificar el trabajo realizado, continuado con la donación, solicitando modificaciones o definitivamente negándose a continuar con la donación. En caso de aceptar seguir con el proceso de donación, el material se guarda en los servidores asignados para ello y todo ello se sigue mediante un contrato inteligente que registra las acciones en la blockchain.

Fig. 1 secuencia de acciones desde el ingreso de los datos clínicos del donante hasta su almacenamiento definitivo.

El proceso de donación conlleva varias acciones como se puede ver en la Fig. 1, y que serán registradas según vayan ejecutando, en una base de datos SQL (BDSQL). Veremos en detalle los dos procesos de anonimización considerados: anonimización automática y manual; este último difiere del primero en que el proceso de eliminar datos personales lo realiza un humano, en lugar de tratarse de una tarea ejecutada por un código que identificaremos como contrato inteligente y que reside en la blockchain.

Proceso de anonimización automática

La imagen en formato DICOM (por sus siglas en inglés, Imaging and Communications in Medicine), permite una identificación univoca de objetos. Cada fichero DICOM tiene un UID único compuesto por varios números. El formato DICOM cuenta con dos objetos: 1) IOD (por sus siglas en inglés, Information Object

Definition) formados por la imagen y su información asociada que juntos forman una representación lógica de objetos del mundo real, 2) DIMSE (por sus siglas en inglés, DICOM Message Service Element) que son operaciones que pueden realizarse sobre un objeto.

IOD y DICOM forman SOP, la unidad funcional de DICOM. Un IOD se compone de las IE (Entidades de información). Hay las IE de paciente, de estudio, de serie, de equipo, de imagen, etc., que a su vez se componen de uno o varios módulos, y que estos módulos a su vez contienen varios atributos. Un atributo en un objeto DICOM, se define con nombre, etiqueta, tipo y descripción; por lo que es factible modificar mediante código las etiquetas originales.

En nuestro caso, lo anterior nos permite borrar la información que identifique al propietario de los datos, esto es, anonimizar el objeto DICOM (ver un ejemplo de esta tarea en Fig. 2).

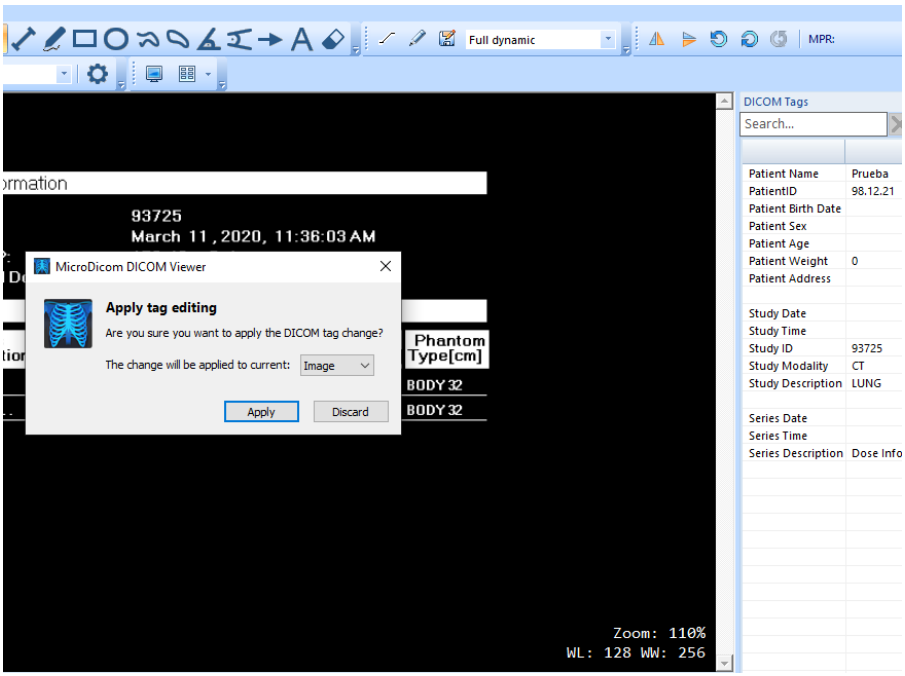


Fig. 1 Modificando el nombre del paciente en un archivo DICOM

La tarea antes descrita la realiza un contrato inteligente contenido en la blockchain, en cada uno de los campos que relacionan al propietario de los datos de manera directa o indirecta (anonimización de la imagen). La siguiente acción es el registro de hallazgos y el diagnóstico asociado a la imagen, tarea a cargo de personal del área médica, quien encripta su nombre con la llave pública del sistema y con este dato firma su actividad, esto implica que ante un proceso de auditoria es posible conocer quien fue el responsable del diagnóstico. Como tarea intermedia cada una de las acciones ejecutadas sobre el objeto DICOM se registran en una base de datos SQL cuya estructura se muestra en la Fig. 3 y se describe el significado de cada uno de los identificadores en la Tabla 1.

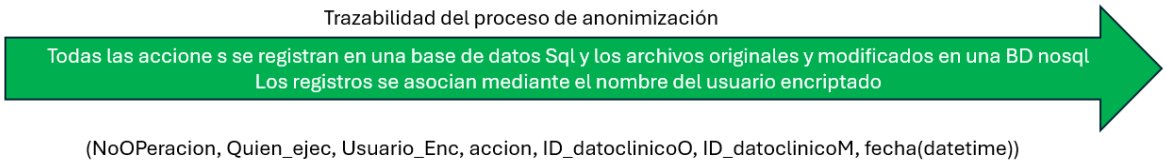


Fig. 2 Ejemplo de registro en la base de datos SQL y NoSQL

Se cuenta con una base de datos NoSQL, en ella se almacenan los objetos DICOM originales y modificados. La intención de contar con ambos archivos es mantener evidencia de referencia sobre la tarea de anonimización que estamos ejecutando. Así el donador puede dar seguimiento al proceso y concluir sin o con la menor cantidad de dudas respecto a la tarea de anonimización. La siguiente tarea es el almacenamiento del objeto DICOM modificado en el almacén de datos. Cumplida esta tarea, todos los registros en la base de datos SQL relacionados con la acción de donación generaran un bloque que se agregara a la blockchain de donaciones. A la información en este bloque tendrá acceso el donador mediante un código QR que se le envía a su correo. En este el punto el donador perderá toda referencia con sus archivos, ya que toda la información relacionada con el proceso de donación es eliminada de la BDSQL al igual que de la BDNoSQL (Base de datos NoSQL).

Tabla 1 Descripción de los campos en la base de datos SQL

Característica	Descripción
NoOperacion	Identificador de una acción de donación
Quien_ejec	Identificador único de quien ejecuta la tarea SHA256
Usuario_Enc	Nombre del donador encriptado con su llave privada
Acción	Que información del archivo DICOM se elimino
ID_datoclinicoO	Identificador generado en la biblioteca del objeto a donar
ID_datoclinicoM	Identificador generado en la biblioteca del objeto procesado
fecha	Fecha con día y hora de la acción ejecutada

Proceso de anonimización manual

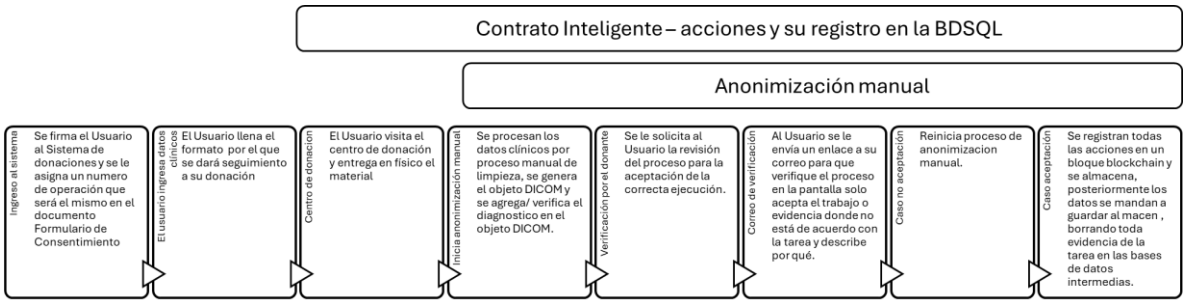


Fig. 4 Ssecuencia de acciones desde el ingreso de los datos clínicos del donante hasta su almacenamiento definitivo.

En este proceso los datos clínicos no están encapsulados en un objeto DICOM, aquí tenemos documentos en físico que contienen los datos clinicos. Así pues, cualquier información que relacione al propietario de los datos clínicos de manera directa o indirecta no es considerada durante el proceso de digitalización y transformación a formato DICOM. Como en el caso anterior, se agrega los hallazgos y el diagnóstico al objeto DICOM y es firmada la acción por el médico a cargo, que encriptará su nombre con la llave pública del sistema, que en caso de auditoria se podrá conocer quien realizó la tarea. En el proceso de anonimización manual, la interacción entre el donador y el humano que procesa los datos es muy relevante para el éxito de la actividad. Cuando existen acciones en el proceso de anonimización, en las cuales el donador no está de acuerdo se sugiere interactuar con el donante para conocer de primera mano su desacuerdo.

Dos condiciones se pueden dar aquí, que el donante no esté de acuerdo en continuar el proceso de donación, aun cuando ya firmó un documento de cesión de sus datos, o que no tenga ningún inconveniente. El primer caso todos los datos son borrados, y se revoca el documento de Formulario de Consentimiento, finalmente se entrega el un código QR a través del correo electrónico del donante para que verifique la tarea. Por otro lado, si no tiene inconveniente, se envía el documento QR con la traza del proceso con un paso adicional, el guardado de los datos en el almacén.

Desarrollo del proceso de certificación de la anonimización de datos clínicos

Lo hasta ahora descrito, considera el desarrollo de un sistema computacional. En los siguientes párrafos se describe la evolución del proceso en el corto y mediano plazo.

Corto Plazo (6 meses a 1 año):

- **Desarrollo Inicial:**

Esta etapa comprende el estudio de la legislación para un proceso de donación de material clínico en México, el trabajo se realizó con el apoyo de abogados con experiencia en el manejo de este tipo de datos.

- **Implementación Temprana:**

Primero se generó la documentación relacionada con el proceso de donación y el algoritmo del proceso de anonimización de la información clínica, donde las fuentes que se consideran son archivos físicos que el propietario de los datos lleva directamente a los centros de acopio, y que posterior a esta entrega, todo el flujo de actividades se realiza en el por medio del micrositio web de donaciones. De forma electrónica, el propietario transfiere archivos digitales y todo el proceso de donación

se realiza desde el micrositio web, donde carga todos los documentos requeridos y se le comparte información que asegura el cumplimiento de la normatividad legal.

- **Adopción Inicial:** En esta etapa el proceso de donación se realiza en el micrositio de donaciones que se encuentre en línea (http://bmdm.citedi.mx/). Su difusión se realiza en las universidades donde tenemos cobertura y en las redes sociales de las mismas. Una tarea actual es la búsqueda de colaboración con asociaciones filantrópicas de las áreas médicas, IMSS-BIENESTAR HIDALGO y sus centros de investigación.

Mediano Plazo (1 a 3 años):

- **Escalabilidad:** Dado que la herramienta es parte de un ecosistema, una de nuestras intenciones es conectarlo a laboratorios y hospitales privados. Se considera el sector salud de México del estado de Hidalgo. Lo anterior implica modificaciones que permitan la interoperabilidad con almacenes de datos o sistemas generadores de datos, que fuera de su interés integrarse. Por lo que se debe buscar en conjunto el equipo legal y el equipo técnico: Como automatizar el proceso de donación de datos clínicos por entidades morales de salud de tal forma que se cumpliera a plenitud con la legislación actual.

- **Mejoras y Actualizaciones:** Se debe primero estudiar la evolución de la legislación vigente en materia de datos clínicos, lo anterior nos permitirá sentar las bases del proceso de actualización del sistema. A mitad de esta etapa es importante realizar un análisis de usabilidad para las mejoras del sitio. El backend consideramos que se debe actualizar casi al final de esta etapa.

- **Expansión del Mercado:** En este momento se prevé que el sitio de la Biblioteca Digital Mexicana de Datos Médicos cuente con la suficiente difusión para convertirse en un espacio seguro para el acceso a datos y capacitación.

Precisar quienes serán los titulares que podrán beneficiarse con la buena práctica o elemento innovador propuesto

Si bien el proceso de trazabilidad su tarea principal es darle certidumbre a los propietarios de los datos clínicos de cómo se protegen sus datos personales e información sensible, el impacto permea a los usuarios de forma directa o indirecta de los datos donados, como se describe a continuación:

- En primera instancia beneficia **a los propietarios de los datos clínicos** que desean hacer donaciones a la biblioteca digital, ya que el proceso de

trazabilidad les garantiza conocimiento pleno de lo que se está haciendo con sus datos personales e información sensible dentro del ecosistema de datos, así el cumplimiento de lo estipulado en el formulario de consentimiento, Anexo A.

- Segundo, dado que una mayor calidad en los datos generada por procesos certificados desde su ingreso por el donante hasta su almacenamiento abona al éxito de los proyectos que los usen, para el caso del ecosistema de datos donde se aplicara este mecanismo, esta mejora en los datos impactara en la **salud de la población mexicana**, cuando los datos se usen con esa intención.
- **Y por último los usuarios de los datos en los sectores médicos privado y público**, así como la academia se beneficia de mejores datos para el entrenamiento de modelos de Inteligencia artificial para sus laboratorios virtuales de capacitación, derivado de la calidad de los datos recolectados.

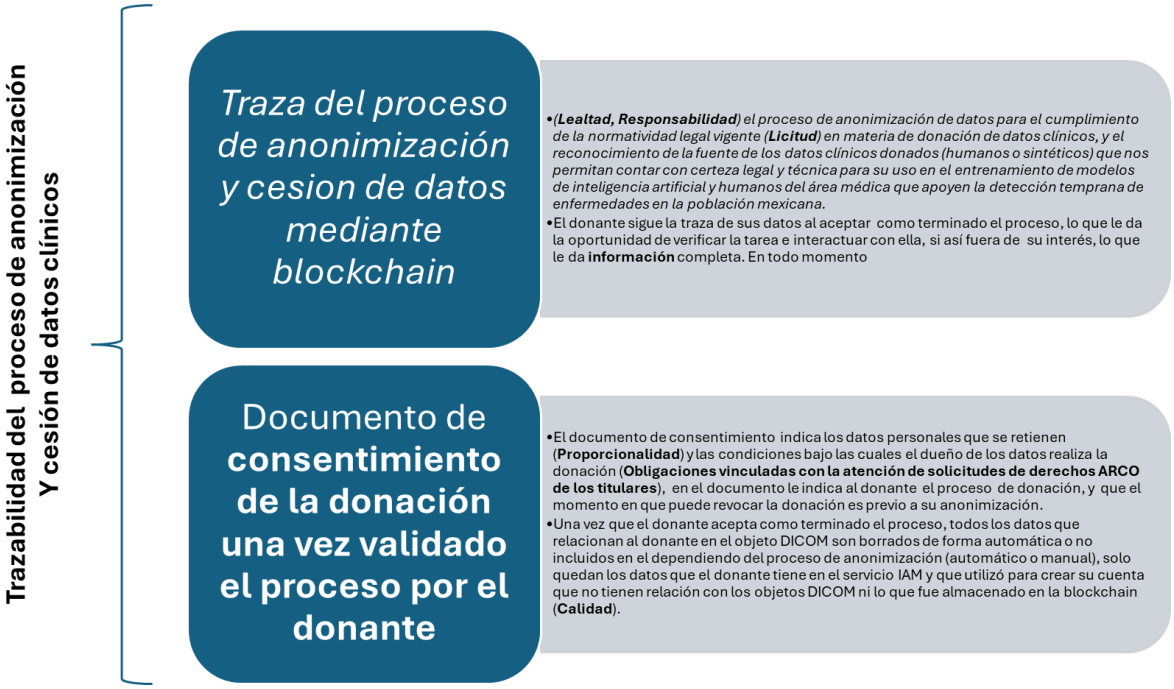
Describir los principales resultados o beneficios que se esperan obtener en caso de que el proyecto sea implementación por un responsable o encargado del sector público o privado.

La trazabilidad del proceso fomenta la transparencia, el acceso a la información y/o la rendición de cuentas respecto a la aplicación cabal de la legislación vigente en materia de datos personales, esto aumenta su importancia cuando la información proviene de grupos vulnerables, como el caso de menores de edad.

Permiten contar con certeza legal y técnica para el uso de datos clínicos en el entrenamiento de modelos de inteligencia artificial que apoyen la detección de enfermedades en la población mexicana, el entrenamiento de personal o estudiantes del área médica y el desarrollo de aplicaciones que requieran el uso de este tipo de datos.

Con la trazabilidad en el blockchain del trabajo realizado (POW), se puede convertir en evidencia y ello permitirá dar valor a los objetos DICOM contenidos en el almacén de datos. Por ende, los servicios prestados por el ecosistema podrían ser monetizados en beneficio del propietario del ecosistema, si este es privado y en beneficio del ecosistema si es público.

Desarrollo de un cuadro sinóptico en el cual se relacione la innovación o buena práctica propuesta con aquellos principios, deberes u obligaciones que sean considerados en el alcance del proyecto.



Definiciones asociadas al cuadro sinóptico.

- **Principios** en materia de protección de datos personales:
  - Licitud: obliga al responsable a tratar los datos personales con apego a lo dispuesto por la normativa aplicable y, en el sector público, de conformidad con sus facultades o atribuciones.
  - Lealtad: impone la obligación al responsable a tratar los datos personales privilegiando la protección de los intereses del titular y la expectativa razonable de privacidad, así como a no utilizar medios engañosos o fraudulentos para recabar y tratar los datos personales.
  - Consentimiento: prevé que el responsable de recabe el consentimiento del titular para el tratamiento de sus datos, salvo las excepciones previstas por la ley.
  - Información: establece que el responsable tiene que comunicar al titular sobre la existencia y características principales del tratamiento al que someterá los datos personales, a través del aviso de privacidad.
  - Proporcionalidad: prevé la obligación a cargo del responsable a tratar únicamente los datos personales que sean necesarios, adecuados y relevantes en relación con las finalidades para las que se hayan obtenido.

- Finalidad: establece la obligación al responsable de tratar los datos personales para las finalidades (concretas, lícitas, explícitas y legítimas) para las cuales se obtuvieron los datos personales y fueron informadas al titular, y consentidas por este último, en su caso.
- Calidad: obliga al responsable a tomar las medidas necesarias para procurar que los datos personales tratados sean pertinentes, correctos, exactos, completos y actualizados, y que se eliminen una vez que concluyó el tratamiento y los plazos de conservación respectivos.
- Responsabilidad: establece la obligación al responsable de velar y acreditar el cumplimiento de los principios antes señalados, deberes y obligaciones, y adoptar medidas para su aplicación, así como de rendir cuentas por los datos personales que trata él mismo o los encargados que los tratan a su nombre y por su cuenta.

- **Obligaciones vinculadas con la atención de solicitudes de derechos ARCO de los titulares,** es decir, para: i) acceder a la información que les pertenece, ii) rectificar o corregir la información que no sea exacta o no esté actualizada, iii) cancelar o eliminar los datos personales, cuando proceda; iv) oponerse a determinados tratamientos de sus datos personales, **así como para revocar el consentimiento que, en su momento, se hubiere otorgado para el tratamiento de sus datos personales**.

Descripción clara de los elementos que permitan determinar la forma en que el proyecto puede ser replicado por los responsables y/o encargados de los sectores privado o público.

Para el desarrollo de la propuesta que aquí se describe se requiere el uso de servicios IAM (gestión de acceso e identidad), tecnología blockchain, manejadores de bases de datos SQL y NoSQL y programa para la generación de códigos QR para mostrar al donante la traza de su proceso.

Servicios IAM (gestión de acceso e identidad)

Dado el tipo de datos que se manejan la mejora de la seguridad es relevante, así reducir el riesgo de accesos no autorizados y ataques cibernéticos mediante la implementación de políticas de seguridad rigurosas es fundamental. Este tipo de servicios facilita el cumplimiento de regulaciones y normas de seguridad al proporcionar registros y auditorías detalladas de accesos y actividades.

Para nuestro caso el acceso al servicio de donación contempla un formato que se anexa a este documento (Formulario de Consentimiento, Anexo A), este formato se firma digitalmente por el donador y ahí mismo se indica que: *No se realizarán transferencias o relaciones de datos personales que permitan la identificación*

individual de cada participante y su material en la biblioteca de imágenes médicas y expedientes clínicos.

Condición que se ratifica al permitir al usuario participar de forma activa en el seguimiento de la anonimización de sus objetos DICOM, y como se menciona antes, si fuera necesaria realizar una auditoría del seguimiento por parte del donador u otra entidad con capacidad legal para hacerlo.

Tecnología blockchain

Esta tecnología por otro lado permite transparentar las acciones que se realizan en la plataforma entorno a la anonimización de objetos DICOM y que finalmente permitirá proveer al usuario de esta información en el momento que lo requiera.

Gestor de base de datos SQL

Este gestor permite el registro de las acciones realizadas por las personas y los contratos inteligentes (código ejecutable) a cargo del proceso de anonimización y diagnóstico del material clínico. También es la fuente de información temporal para la generación del bloque que se agregara a la blockchain.

Gestor de base de datos NoSQL

La tarea de este elemento es el almacenamiento de los datos no estructurados originales, objetos DICOM y sus modificaciones durante el proceso de anonimización y registro de diagnóstico. Es la fuente de información temporal para el almacén de datos.

Generador de llaves asimétricas

Estas llaves nos permiten opacar al donador de las imágenes, a quien las procesa y diagnostica, durante todo el proceso de anonimización y en su registro a la blockchain.

Generador de códigos QR.

Le da acceso al donante a la información almacenada en el blockchain relacionada con su proceso de donación y a los elementos de código que apoyaron la tarea.