ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA DE SISTEMAS

IMPLEMENTACIÓN DE MECANISMOS DE CONTROL DE ACCESO A INFORMACIÓN DE IDENTIFICACIÓN PERSONAL (PII) DE ACUERDO CON LA LEY ORGÁNICA DE PROTECCIÓN DE DATOS PERSONALES DEL ECUADOR

MECANISMOS PARA LA RECOLECCIÓN Y PROCESAMIENTO SEGURO DE INFORMACIÓN DE IDENTIFICACIÓN PERSONAL (PII)

TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN CIENCIAS DE LA COMPUTACIÓN

AUTOR: WILLIAM DARIO SUNTAXI PICHUASAMIN

willian.suntaxi@epn.edu.ec

DIRECTOR: DENYS ALBERTO FLORES ARMAS

denys.flores@epn.edu.ec

DMQ, marzo 2023.

CERTIFICACIONES

Yo, WILLIAM DARIO SUNTAXI PICHUASAMIN declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

WILLIAM DARIO SUNTAXI PICHUASAMIN

Certifico que el presente trabajo de integración curricular fue desarrollado por WILLIAM DARIO SUNTAXI PICHUASAMIN, bajo mi supervisión.

DENYS ALBERTO FLORES ARMAS

Certificamos que revisamos el presente trabajo de integración curricular.

NOMBRE_REVISOR1	**NOMBRE_REVISOR2**
REVISOR1 DEL TRABAJO DE INTEGRACIÓN CURRICULAR	**REVISOR2 DEL TRABAJO DE INTEGRACIÓN CURRICULAR**

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el (los) producto(s) resultante(s) del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

Adicionalmente, declaro que el contenido de este trabajo incluye material de artículos técnicos revisados por pares de mi autoría que han sido publicados durante el desarrollo de este proyecto, y que detallo a continuación:

Firmado electrónicamente por:
WILLIAM DARIO
SUNTAXI PICHUASAMIN

WILLIAM DARIO SUNTAXI PICHUASAMIN

Firmado electrónicamente por:
DENYS ALBERTO
FLORES ARMAS

DENYS ALBERTO FLORES ARMAS

DEDICATORIA

Esta tesis está dedicada a:

A la reina de mi vida, mi hermosa madre Clemencia y a mi padre J. Aníbal quienes han sido el motor, inspiración, ejemplo, y eje de sabiduría durante toda mi vida académica. Gracias por todas las charlas y consejos de madrugada. Gracias por nunca dudar de mí.

A mis hermanos Jorge y Carina por su apoyo incondicional. Gracias a sus consejos y enseñanzas se ha logrado este triunfo. A toda mi familia porque gracias a sus oraciones, consejos y cuidado han logrado que este suceda.

Como olvidar a mi sobrino Bruce, a mis sobrinas Gabriela, Antonella e Isabella que, con su inocencia y amor, han logrado convertirme en un mejor ser humano. Me enseñaron el amor bonito.

A mi cuñado Juan Carlos y a mi cuñada Anita que me han permitido entrar a sus vidas para tener una relación de amistad, lealtad y sinceridad ante cualquier adversidad.

Finalmente quiero dedicar esta tesis a todos mis amigos, que siempre los he considerado como mis hermanos. Agradezco su amistad sincera y real, gracias por estar en momentos alegres y difíciles de mi vida.

William D. Suntaxi P.

AGRADECIMIENTO

Al finalizar este trabajo quiero utilizar este espacio para agradecer a la reina de mi vida, a mi madre Clemencia que siempre me ha dado su ejemplo de lucha, honestidad, paciencia y sobre todo de trabajo. Ante cualquier eventualidad, enfermedad o distancia ella nunca ha dejado de ser la madre, esposa y abuelita amorosa. Gracias por enseñarme a ser feliz en la vida amor de mi vida, este logro sin duda alguna es para ti.

A mi padre:

Aníbal Suntaxi Guayasamín quién con su ejemplo y amor ha logrado convertirme en un hombre de bien, gracias por acompañarme y darme todo desde pequeño para poder ser un profesional en la vida.

A mi hermano:

Jorge Suntaxi quien me enseño todo lo que se ahora, gracias por ser un hermano amoroso, gracias por compartir todo tu aprendizaje hermano de mi corazón. Te agradezco por inculcar en mí, todo tu conocimiento.

A mi hermana:

Carina Suntaxi quien se preocupa por todo lo que sucede a mi alrededor, por ser el soporte y la unión de la familia. Te amo demasiado reina, gracias por el aguante durante este proceso de aprendizaje.

A mis amigos:

Rafael Rivera y Christian Lagua, mis amigos de la adolescencia. Los dos se han convertido en mis hermanos con el pasar de los años, les agradezco por toda su ayuda, respeto y amistad real en todo el sentido de la palabra. Se logró hermanos, este también es un triunfo de ustedes.

A Roberto Toapanta, Danilo Angamarca, José Valencia, Erick Rivera, y a todos los que conforman mi equipo Full Norton. Agradezco por su amistad sincera durante la carrera en la Universidad. Por todos los momentos compartidos y por los cuales somos quienes somos en la actualidad, les quedaré eternamente agradecido.

William D. Suntaxi P.

ÍNDICE DE CONTENIDO

RESUMEN

La presente investigación realiza la implementación de un aplicativo web para la recolección de información de manera segura, siguiendo los lineamientos establecidos en la Ley Orgánica de Protección de Datos Personales LOPDP. Mediante la pseudonimización y la minimización de los datos personales se busca reducir la filtración de datos personales. Una vez identificados los datos que deben ser tratados se realiza la persistencia de información en una base de datos intermedia que permite el almacenamiento de datos minimizados y pseudonimizados para posteriormente mostrarse como datos públicos a una entidad externa. Para mostrar la información privada de un usuario la entidad externa tiene que pasar por una capa de seguridad en donde se verifica y autentica utilizando certificados digitales y la verificación con doble factor utilizando un OTP. Posteriormente se realiza las pruebas de rendimiento para verificar como los recursos y funcionalidad del aplicativo se ven afectados con el aumento de seguridad. Para las pruebas de rendimiento se realizaron sobre 3 componentes principales, disco duro, memoria RAM y memoria caché en los cuales se obtuvieron resultados favorables. Finalmente, se realizó un modelo de riesgos y amenazas para identificar las vulnerabilidades y buscar una forma de mitigarlo sin comprometer el rendimiento y la disponibilidad del aplicativo web.

PALABRAS CLAVE: minimización, pseudonimización, certificados digitales, OTP.

ABSTRACT

In this research, the implementation of a web application for the collection of information in a secure manner, following the guidelines established in the Organic Law on Personal Data Protection LOPDP is carried out.

Using pseudonymization and minimization of personal data, the aim is to reduce the leakage of personal data. Once the data to be processed has been identified, the information is persisted in an intermediate database that allows the storage of minimized and pseudonymized data to be subsequently shown as public data to an external entity.

To show a user's private information, the external entity has to go through a security layer where it is verified and authenticated using digital certificates and two-factor verification using an OTP.

Subsequently, performance tests are performed to verify how the increased security affects the application's resources and functionality.

Performance tests were performed on 3 main components, hard disk, RAM, and cache memory in which favorable results were obtained. Finally, a risk and threat model was performed to identify vulnerabilities and find a way to mitigate them without compromising the performance and availability of the web application.

KEYWORDS: minimization, pseudonymization, digital certificates, OTP.

1. INTRODUCCIÓN

En este capítulo se pretende abordar y describir de forma generalizada aspectos del componente, objetivo general y objetivos específicos, alcance del proyecto, fundamentación teórica y finalmente los trabajos previos relacionados del proyecto a desarrollar.

1.1 Descripción del Componente

En la actualidad, existe en la gran mayoría de las organizaciones privadas o públicas en el Ecuador un predominante desconocimiento de mecanismos seguros para la gestión de los datos personales de los individuos. Esto se debe esencialmente al desconocimiento de las leyes estatales, como la Ley Orgánica de Protección de Datos Personales del Ecuador (LOPDP) por ejemplo, la cual contiene dentro de ella una serie de artículos que otorgan una idea global de como deberían tratarse y procesarse los datos personales de los individuos que radican dentro del territorio ecuatoriano. A su vez, los datos personales recolectados por cada organización deberán cumplir con ciertos lineamientos establecidos en la LOPDP, uno de ellos es el principio de Pertinencia y Minimización, los cuales deben estar limitados a lo estrictamente necesario para el cumplimiento de la finalidad del tratamiento. Cabe recalcar que la entrega de los datos personales por parte del titular es realizada bajo su previo consentimiento; el titular debe conocer a plenitud la finalidad con la cual serán procesados sus datos personales. La finalidad del tratamiento hace hincapié a los datos necesarios que requiere la organización como cumplimiento de su objetivo de negocio. Por otra parte, la LOPDP menciona, dentro de su ley, un equipo de trabajo que forma parte del sistema de protección de datos personales; este equipo está conformado por el responsable del tratamiento, el encargado del tratamiento y el delegado de protección de datos personales. Este equipo tendrá entre sus principales funciones recolectar, procesar y almacenar los datos personales de los titulares dentro de la organización. El rol del responsable está en dar a conocer de forma clara y precisa al titular de los datos para que serán tratados sus datos y cuáles son todos sus derechos en ese ámbito. Los roles del delegado y encargado del tratamiento será un trabajo en conjunto con aspectos técnicos y de control interno para la protección de los datos personales almacenados. No obstante, la LOPDP se limita a indicar como deben implementarse de forma técnica sus artículos por parte de quienes participan dentro de la ley, ni tampoco menciona aspectos de herramientas computacionales y servicios que faciliten el proceso del tratamiento de la información recolectada. Por consiguiente, surge la necesidad de proponer e implementar mecanismos técnicos de gestión de la información que guíen al equipo de sistema de protección de datos personales en cumplimiento con la LOPDP.

Estos mecanismos serán de gran apoyo para el equipo ya mencionado de la organización, evitando errores durante el proceso de tratamiento y sanciones posteriores a la organización por parte de la Autoridad de Protección de Datos Personales.

1.2 Objetivo general

Implementar mecanismos seguros de gestión de la información en el ámbito de la identificación personal (PII), mediante un adecuado control de acceso a la misma durante su recolección y procesamiento, garantizando la seguridad de la PII.

1.3 Objetivos específicos

Establecer al menos tres objetivos específicos. Los objetivos específicos detallan los procesos necesarios para la completa realización del componente; sirven como una guía de la manera en la que será abordado el componente asignado.

1. Analizar el estado del arte sobre las soluciones para la recolección y procesamiento de PII.
2. Identificar los requerimientos funcionales y no funcionales relacionados con la seguridad y privacidad de PII durante su recolección y procesamiento.
3. Implementar un prototipo considerando como referencia el GDPR europeo, la LOPDP ecuatoriana y el Esquema Gubernamental de Seguridad de la Información (EGSI).
4. Analizar los resultados de evaluación del prototipo en un entorno controlado.

1.4 Alcance

Se contempla la implementación de mecanismos de control de acceso durante su recolección y procesamiento, garantizando la seguridad de la PII respetando el derecho a la privacidad de los propietarios o custodios de la información.

Se considerará dentro de este proyecto como referencia el RGPD europeo, la LOPDP ecuatoriana y el Esquema Gubernamental de Seguridad de la Información (EGSI) para definir requerimientos funcionales y no funcionales.

Los mecanismos implementados son prototipos experimentales y serán evaluados en entornos controlados, sin que esto implique el despliegue de estos en alguna infraestructura productiva, ya sea pública o privada. Para ejemplificar el correcto proceso de consentimiento, entrega y procesamiento de los datos personales, se propone un esquema que se muestra en la siguiente Figura 1.

Figura 1: Proceso de recolección, procesamiento de datos personales.
[Autor: William Suntaxi]

Como se puede observar en la Figura 1, se plantea esta solución para pseudonimizar los datos personales de los individuos que otorguen el consentimiento para el tratamiento de estos. En concordancia con la Figura 1, se puede describir cada etapa en 6 pasos que se describen a continuación:

1) Datos de identificación personal: Son todos los datos que identifican a un sujeto, entre estos pueden ser su número de cédula, número telefónico o dirección domiciliaria.
2) Consentimiento: Documento legal el cual es firmado por el titular de los datos personales, el cual otorga la autorización a que se puedan tratar o procesar sus datos para un fin específico y durante un tiempo determinado.
3) Inicio de minimización: Se ejecuta la estrategia de minimización durante la recolección de datos; la principal tarea de la minimización es tratar una pequeña cantidad de datos personales con el fin de evitar el tratamiento de datos innecesarios.
4) Patrón de diseño: se elige el patrón con el cual se minimizarán los datos, por ejemplo, la técnica de pseudonimización.
5) Técnica de pseudonimización: técnica que forma parte de la estrategia de la minimización, esta podría ser, por ejemplo, la técnica de la clave secreta.
6) Datos pseudonimizados: Son los datos que han pasado por todo el proceso de minimización.

1.5 Marco teórico

Para el estudio de mecanismos para recolección y procesamiento seguro de Información de Identificación Personal (PII), se considerarán los siguientes fundamentos teóricos, los cuales permitirán dar un contexto claro y conciso del estudio en curso sobre las leyes actuales en el marco de la protección de los datos personales.

1.5.1. Ley Orgánica de Protección de Datos Personales

La Asamblea Nacional del Ecuador, en el Registro Oficial Suplemento No. 459 del 26 de mayo de 2021, la Función Legislativa promulgó la primera ley referente a la protección de los datos personales y que es titulada como Ley Orgánica de Protección de Datos Personales (LOPDP) [1]. Esta Ley tiene como principal objetivo garantizar el derecho a la protección de datos personales de los ecuatorianos; esto circunscribe el acceso, decisión y protección a la información y a los datos de esta índole.

La presente Ley está dividida en 12 capítulos, en los cuales se abordan los principios, derechos, categorías especiales, transferencia o comunicación, seguridad y protección, de los datos personales. Estos capítulos se ven respaldados a través de 77 artículos que van manifestando como y porque está formada la Ley, de cómo debe ser entendida y ejercida y de las partes que actúan en el marco legal [1].

Por lo que se refiere al ámbito de aplicación material (Artículo 2 de la LOPDP), el tratamiento o procesamiento de datos personales será aplicado en cualquier tipo de soporte ya sea automatizado o no, y que el tratamiento empezará cuando el dueño de los datos haya otorgado su consentimiento. Una vez otorgado el consentimiento, se procederá a la recolección [2].

Al mismo tiempo, la LOPDP está encargada de enunciar los diferentes contenidos referentes a los derechos de los titulares de datos personales en espacios referentes a:

1. Información personal;
2. Acceso a datos del titular;
3. Derecho a la corrección y actualización de datos personales del titular;
4. Derecho a la eliminación de datos personales;
5. Derecho a la oposición;
6. Derecho a la portabilidad y transferencia;
7. Derecho a la no decisión fundada en valoraciones automatizadas;

8. Derecho a la consulta pública y sin costo dentro del Registro Nacional de Protección de Datos Personales;
9. Derecho a la educación digital.

Asimismo, la LODP establece un tratamiento especializado categorizando los datos personales según diferentes condiciones, entre estos se mencionan los datos sensibles, de salud, de niños, niñas y adolescentes, y datos de personas con discapacidad o capacidades especiales [2].

Finalmente, la ley consagra un régimen de sanciones para infracciones relacionadas con el tratamiento de datos personales. Las normas relativas al régimen sancionatorio no entrarán en vigor sino hasta dos años después de la publicación de la ley en el Registro Oficial [1].

1.5.2. Reglamento general de protección de datos

El RGPD establece los requisitos específicos para empresas y organizaciones sobre la recopilación, almacenamiento y gestión de los datos personales. Se aplican tanto a las organizaciones europeas que tratan datos personales de ciudadanos en la UE (Unión Europea) como a las organizaciones que tienen su sede fuera de la UE y cuya actividad se dirige a personas que viven en la UE [3], [4].

El RGPD entra en funciones de la siguiente manera:

- Las organizaciones dentro de la UE procesan y recopilan datos personales en dichos territorios.
- La organización tiene una locación externamente de la UE, pero que recopila y trata datos personales de ciudadanos dentro de la UE.

Finalmente, las organizaciones que no residan dentro de la UE y cuyo propósito de trabajo sea la recolección de datos de ciudadanos de la UE mandatoriamente deben nombrar un representante dentro de la UE [5].

1.5.3. Esquema Gubernamental de Seguridad de la Información

El Esquema de seguridad de la información del gobierno (EGSI) tiene como objetivo proteger la privacidad, la integridad y la disponibilidad de la información mediante el uso de métodos de gestión de riesgos de la información y la selección de controles de gestión de riesgos [6].

1.5.4. Otras Leyes Referentes a la Protección de Datos Personales

Los países buscan esforzarse al establecer políticas estatales de protección de datos. Según la consultora Gartner Inc. [7] el 65% de la población mundial tendrá sus datos personales protegidos por una legislación moderna de privacidad en el 2023.

En 2023, entrará en vigor en el estado de Virginia (EE. UU.) la ley conocida como la CDPA (Consumer Data Protección Act) [8]. Esta ley menciona como las organizaciones deberán obtener permisos para el procesamiento de datos personales y al mismo tiempo permitir a los residentes escoger por no participar si van a ceder sus datos con fines económicos.

Por otra parte, en septiembre del 2021, entró en vigor la Ley de Protección de la Información Personal de China [9] con su primera tentativa de instaurar una normativa sobre la privacidad de los datos en ese país.

A finales de 2020, en Brasil, se propone la Ley General de Protección de Datos (LGPD) [10], el cual permite proteger los datos personales de aproximadamente 140 millones de usuarios en América Latina.

1.5.5. Minimizar

La minimización es una estrategia de diseño que comienza durante la recolección de datos; el principal cometido de la minimización es tratar una pequeña cantidad de datos personales con el fin de evitar que el tratamiento de datos innecesarios debido a los fines perseguidos durante el tratamiento [11]. Además, esto puede conseguirse al recolectar datos de menos sujetos (reducir el tamaño de la población de estudio) o menos datos de los sujetos (reducir el volumen de información recopilada) para lo cual pueden utilizarse las siguientes tácticas:

- **Seleccionar:** elegir únicamente la muestra de individuos relevante y los atributos necesarios siguiendo una actitud conservadora al establecer el criterio de selección y realizar el tratamiento únicamente sobre los datos que respondan a dicho criterio (lista blanca).
- **Excluir:** es el enfoque inverso al anterior, y consiste en excluir de antemano los sujetos y atributos que resulten irrelevantes para el tratamiento realizado (lista negra). En este caso se debe adoptar una actitud abierta, intentando excluir el máximo posible de registros a menos que pueda justificarse que son absolutamente necesarios para la finalidad perseguida.
- **Podar:** eliminar parcialmente los datos personales tan pronto dejen de ser necesarios lo cual supone determinar de antemano cuál es el periodo de conservación para cada

uno de los datos recogidos y establecer mecanismos automáticos de borrado cuando se cumpla dicho plazo. En el caso de que los datos formen parte de un registro en el que figure más información que sea necesario conservar, el valor de los campos no necesarios puede modificarse a un valor por defecto prefijado.

- **Eliminar:** suprimir por completo los datos personales tan pronto dejen de ser relevantes asegurándose que no es posible su recuperación ni siquiera de las copias de seguridad realizadas.

También es necesario tener en cuenta que sólo se deben comunicarse y compartirse los datos estrictamente necesarios y que, en el caso de tratamientos que infieran nueva información personal, también deben seleccionarse para su exclusión aquellos datos que se generen y no sean necesarios para la finalidad perseguida.

De forma general, se plantea un esquema tal cual se muestra en la Figura 2, el cual que refleja la estrategia de privacidad para el caso de la minimización de los datos.

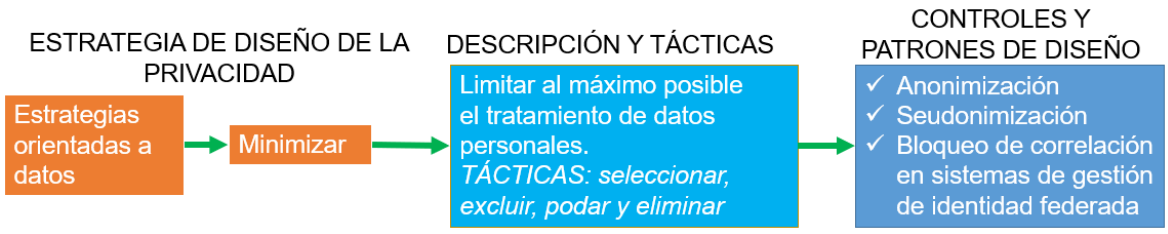

Figura 2: Estrategia de diseño de la privacidad.

[Autor: William Suntaxi]

La estrategia del esquema muestra de inicio a fin, como la minimización de los datos durante la fase de recolección puede utilizar diferentes tácticas que pueden ser escogidas para el diseño de privacidad de los mismos. Luego de seleccionar alguna táctica se elige un patrón de diseño, el cuál va a encaminar el proceso completo de la minimización. Para el actual estudio, se ha planteado el patrón de diseño de minimización llamado pseudonimización.

1.5.6. Pseudonimización

En el RGPD [3] el control o patrón de diseño de la pseudonimización se presenta en información que, sin la necesidad de contener los datos de una persona, permita identificarla por medio de información agregada, siempre y cuando esta sea insertada por separado técnicamente garantizando inminentemente que los datos de identificación personal no sean atribuidos de forma inmediata a una persona física identificable. Este

patrón de diseño trata los datos personales sin tomar en cuenta los datos identificables del titular sin eliminar la vinculación entre datos que logran establecer el sujeto titular de los mismos. Se puede señalar que el RGPD insertó este patrón como una medida lícita para la protección de datos. La Seudonimización es un procedimiento reversible (los datos pseudonimizados conservan datos adicionales que pueden reidentificar a los sujetos).

Algunos ejemplos de pseudonimización de datos se listan a continuación:

- ✓ Cambiar el nombre y apellido de los usuarios por un código alfanumérico que sirva como identificador [12].
- ✓ Cambiar el DNI de los clientes por un código numérico diferente [12].
- ✓ Cuando se recogen muestras biológicas, como el análisis de sangre, se identifican con un código que sustituye al nombre y apellidos del paciente [12].

Para la correcta ejecución de la pseudonimización, es de vital importancia el almacenamiento de la información adicional que permita asignar los datos pseudonimizados a su propietario.

La pseudonimización se puede realizar por cualquier de las siguientes maneras:

- ✓ Mediante la sustitución de cifras y códigos por palabras. Una vez que se ha sustituido, esa información no significa nada y para su correcta lectura se debe estar en posesión de la información adicional [13].
- ✓ A través de la codificación de la información. Esta codificación puede ser retornada con la clave de desencriptación [13].
- ✓ Cuando la información se almacene bajo un número aleatorio que carezca de relación con la información original [13].
- ✓ Existe la sustitución de cifras y códigos por palabras y a la vez se cuenta con una clave para su desencriptación [13].
- ✓ Se intercambia un número aleatorio por un conjunto de datos [13].

Existen técnicas de pseudonimización que pueden ser aplicadas según la necesidad de cada organización. Las técnicas se muestran en la siguiente Figura 3:

Figura 3: Técnicas de Pseudonimización.

1.5.7. Técnicas de Pseudonimización

- Cifrado con clave secreta [14]:

  Se emplea una clave secreta para identificar a los interesados. Es decir, se emplea una clave secreta para cifrar el conjunto de datos, de manera que solo quien posee dicha clave puede descifrar (revertir) la base de datos e identificar a las personas físicas a través de los datos. Para asegurar la seguridad de la base de datos, se deben emplear sistemas de cifrados avanzados y controlar en todo momento quién tiene acceso a la clave de cifrado.

- Función con clave almacenada [14]:

  La función con clave almacenada es un tipo de función hash que emplea una clave secreta como valor de entrada suplementario, de manera que, para acceder a los datos, el responsable de tratamiento debe reproducir la ejecución de la función con el atributo y la clave secreta. Las funciones de hash son algoritmos que pueden crear a partir de una entrada de cualquier tipo (texto, archivo, contraseña) una salida alfanumérica de longitud fija, que representa un resumen de toda la información suministrada.

- Cifrado determinista o función hash con clave de borrado de clave [14]:

  Esta técnica genera un número aleatorio como seudónimo de cada atributo de la base de datos (de cada dato personal) y después borra la tabla de correspondencia. De esta manera se reduce el riesgo de vincular los datos personales del conjunto de datos y los datos personales relativos a la misma persona almacenados en otro conjunto de datos, donde el seudónimo utilizado sea diferente.

- Descomposición en tokens [14]:

  La descomposición de tokens se basa en las técnicas anteriores y consiste en aplicar mecanismos de cifrado unidireccionales o en asignar, mediante una función de índice, un número de secuencia o un número generado aleatoriamente y que no derive matemáticamente de los datos originales.

  Se emplea especialmente en el ámbito financiero, ya que cambia los números de identificación de tarjetas por valores sin utilidad alguna, en caso de que alguien consiguiera acceder de manera fraudulenta a estos datos.

1.6 Estado del Arte y Trabajo Relacionado

Dentro de la problemática que se está abordando en este estudio, es necesario analizar los diferentes trabajos y leyes que tratan y mencionan aspectos de recolección y procesamiento de datos personales. En esta unidad se discutirá sobre los trabajos en cuestión.

1.6.1 Estudios sobre la recolección y procesamiento de datos personales

En el estudio de Bruno Santos [15], se tratan de principios jurídicos y conceptos económicos fácticos para la protección efectiva del consentimiento de los titulares de datos personales en el marco de la Política de Privacidad y la Ley General de Protección de Datos (LGPD) de Brasil. En este contexto se sugiere la reflexión para verificar si el consentimiento de la entrega de los datos personales es un instrumento de real eficacia para la tutela de los sujetos en red. Por otro lado, en el estudio denominado Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma ISO/IEC 27001 [16], se proponen lineamientos para la protección de los datos de las organizaciones por medio de la conducción de proyectos de diagnóstico, para la implementación e implantación de sistemas de seguridad de la información – SGSI alineado con el estándar ISO/IEC 27001 y el sistema de control propuesto en la norma ISO/IEC 27002. Dentro de ese estudio se aplicaron fases de auditoría, análisis y evaluación de riesgos con la aplicación de cuestionarios aplicados a los administradores, clave de seguridad, entrevistas al personal del área informática y usuarios de los sistemas, pruebas de intrusión y testeo que permitieron establecer el diagnóstico de seguridad de esa organización. Posteriormente se aplicó una lista de chequeo que permita verificar la existencia de controles de seguridad en los procesos organizacionales. Como último paso se aplicaron los controles de seguridad para que sean integrados dentro del SGSI que responda a las necesidades de seguridad informática y de la información. A su vez, en el

estudio de Susana Echeverría [14], se aplican controles de la norma ISO/IEC 27002 para la seguridad de la información de la Empresa Eléctrica Quito S.A. guiando a sus empleados de la empresa para la identificación de los niveles de desempeño para la protección de sus activos. Al mismo tiempo, otro estudio realzado en 2021 en Ecuador por parte de Felipe Roldán [17], se analizan los principios del consentimiento del titular y la finalidad del tratamiento de datos personales para determinar qué estándar de protección se podría implementar. Este estudio examinó el modelo normativo europeo y estadounidense para cotejar un posible impacto práctico que se podría generar cuando entre en vigor la ley ecuatoriana LOPDP. Finalmente, en el trabajo de investigación llamado Elaboración de Lineamientos para la implementación de los primeros pasos del sistema de seguridad de la información gubernamental EGSI en las agencias del gobierno central [17], se recomiendan los lineamientos para implementar el actual sistema de seguridad de la información gubernamental EGSI en el gobierno de. el gobierno central del Ecuador que permita a los miembros involucrados seguir un conjunto de pasos para realizar la implantación de esta norma.

1.6.2 Contribución

El presente trabajo pretende contribuir de manera significativa en la implementación de mecanismos de control de acceso durante su recolección y procesamiento de los datos personales. La propuesta de implementación está inspirada en la carencia de guías técnico-prácticas que simplifiquen el trabajo de la creación de sistemas que faciliten el proceso de recolección de datos personales utilizando canales y aplicaciones seguras junto con sistemas de bases de datos igualmente protegidas que inminentemente garanticen la seguridad de la PII respetando los derechos a la privacidad de los titulares de la información.

1.6.3 Método Prisma

El método Prisma es una herramienta utilizada en investigación para la selección y evaluación crítica de artículos científicos. El objetivo principal de este método es identificar y seleccionar artículos relevantes para una revisión sistemática de la literatura.

El método Prisma se compone de los siguientes pasos:

1. Identificar la pregunta de investigación: Esta etapa implica definir claramente la pregunta de investigación y los objetivos de la revisión sistemática.

2. Búsqueda de la literatura: Se lleva a cabo una búsqueda exhaustiva y sistemática de la literatura relevante, utilizando bases de datos electrónicas, registros de ensayos clínicos y otras fuentes relevantes.
3. Selección de estudios: Los artículos se seleccionan en base a criterios de inclusión y exclusión predefinidos, como el tipo de estudio, el diseño, el tamaño de la muestra, la calidad metodológica, entre otros.
4. Evaluación de la calidad: Se realiza una evaluación crítica de la calidad metodológica de los estudios incluidos, utilizando herramientas de evaluación de calidad y riesgo de sesgo.
5. Extracción de datos: Se extraen datos relevantes de los estudios seleccionados y se resumen en tablas o matrices.
6. Análisis y síntesis de los datos: Se realiza un análisis y síntesis de los datos extraídos, utilizando técnicas estadísticas y otras herramientas de análisis.
7. Presentación de los resultados: Se presentan los resultados de la revisión sistemática en un informe completo, que incluye una descripción detallada de los métodos utilizados, los resultados del análisis y las conclusiones.

El método Prisma es una herramienta valiosa para la realización de revisiones sistemáticas de la literatura, ya que proporciona una estructura clara y rigurosa para la selección y evaluación de artículos científicos [18].

1.6.4 Revisión sistemática de la lectura

Se realizo un análisis del estado del arte de soluciones para la recolección y procesamiento seguro de PII. La investigación tiene como propósito analizar los estudios existentes y sus hallazgos con el fin de entender la problemática planteada y encontrar todas las referencias posibles para dar solución a las preguntas de investigación planteada:

- ¿A qué se considera un mecanismo seguro de datos?
- ¿Qué mecanismos de seguridad se han implementado para la recolección y procesamiento de datos?
- ¿Cómo se garantiza la integridad de los datos?

Este documento proporcionó grandes contribuciones para aquellas personas que estén interesados en recolección y procesamiento seguro de datos:

- Se identificó 8 estudios principales relacionados a recolección y procesamiento de datos entre 2016 y el 2022.

- Se presentó un metaanálisis del estado actual de los métodos en los que se puede implementar para recolectar y procesar transferir datos de manera segura.
- Se ha hecho una representación y se produjo pautas para apoyar más trabajar en esta área.

Para el artículo antes mencionado se realizó la búsqueda de información en tres fuentes importantes en el ámbito de la ingeniería las cuales son: JSTOR, Science Direct, Google Scholar. Se identificaron un total de 680 estudios relacionados con las palabras clave filtradas en los motores de búsqueda. Se pudo apreciar una considerable reducción al descartar resultados que no eran artículos académicos, lo que dejo un total de 450 resultados finales. Al filtrar la búsqueda por el año de referencia ese resultado bajo a 100. Al establecer los criterios de inclusión y exclusión el resultado bajó a 45, y, por último, después de leer su título y resumen se volvieron aplicar ambos criterios dejando un resultado de 8 estudios.

En la Figura 4 se puede visualizar el proceso de selección de los datos del proyecto como también las fuentes y los resultados iniciales de la búsqueda.

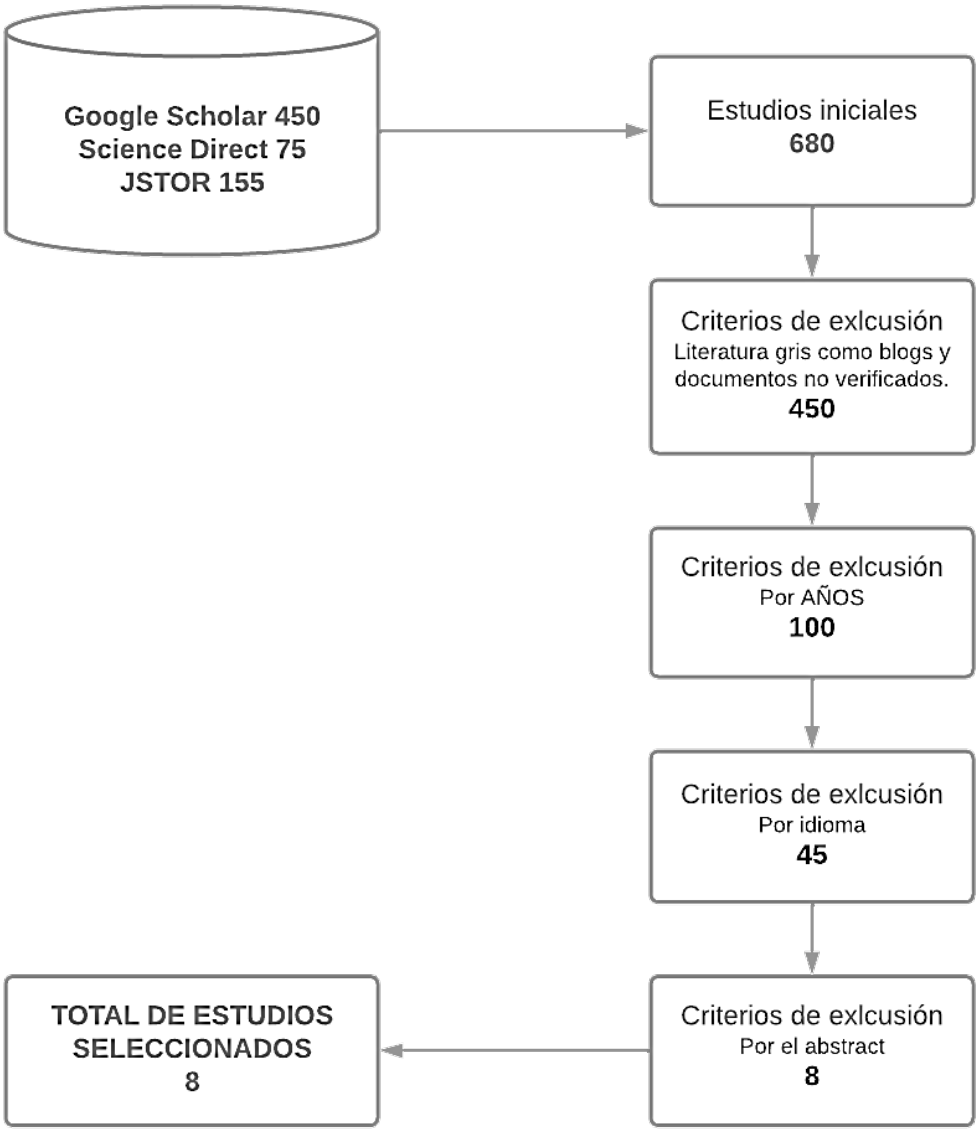

Figura 4: Representación del proceso de selección de información.

[Autor: William Suntaxi]

La Figura 5 muestra el número de estudios primarios publicados cada año. En la figura hay una tendencia al alza en cuanto a la investigación del tratamiento de datos personales. La protección de datos y los mecanismos de recolección de datos. Dado que en Ecuador y en todo el mundo se han creado nuevas leyes sobre la protección, seguridad y tratamiento de datos personales se prevé que en el futuro se verá una cantidad significativa de estudios de investigación relacionados a esta SLR.

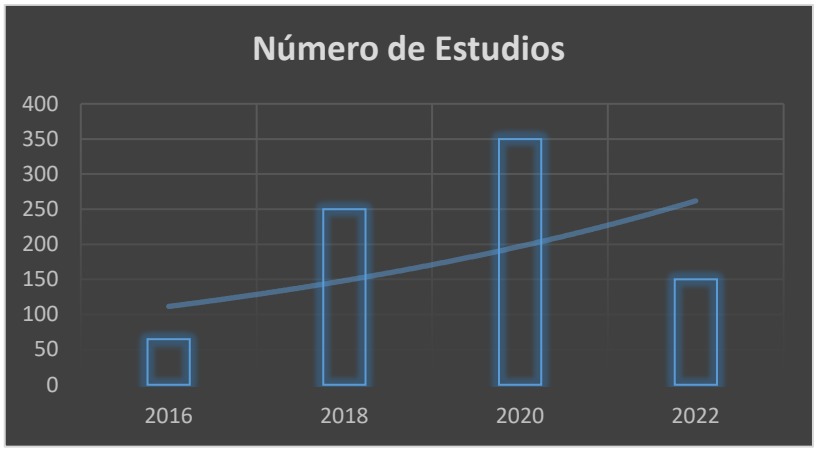

Figura 5: Representación de los trabajos obtenidos filtrado por años.

[Autor: William Suntaxi]

A continuación, se presentan los hallazgos más importantes

Tabla 1. Resultados obtenidos del SLR.

[Autor: William Suntaxi]

No	**Estudios Primarios**		
	Título	**Autores**	**Enfoque**
E1	Reflexiones escépticas, principiológicas y económicas sobre el consentimiento necesario para la recolección y tratamiento de datos.	Santos Divino, S. B.[19]	Recolección y procesamiento de datos
E2	Proyecto Política de la Información	Escuela Politécnica Nacional [20]	Protección de datos personales
E3	Gestión de identidad digital de usuarios en servicios web para la protección de la privacidad de la información (Doctoral dissertation, Ecuador-PUCESE-Escuela de Sistemas y Computación).	Utreras Logacho, P. L. [21]	Protección de datos personales

E4	Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma ISO/IEC 27001. Revista Tecnológica-ESPOL, 28(5).	Solarte, F. N. S., Rosero, E. R. E., & del Carmen Benavides, M. [22]	Protección de datos personales
E5	Modelo de prevención para el tratamiento de datos personales (Doctoral dissertation, Universidad Nacional de La Plata).	Sebastián, M. A., & Vázquez, N. E. [23]	Recolección y procesamiento de datos
E6	Evaluación del sistema de gestión de la seguridad de la información de la Empresa Eléctrica Quito utilizando la norma ISO 27002	Echeverría Rodríguez, Sara Viviana [24]	Recolección y procesamiento de datos
E7	Los ejes centrales de la protección de datos: consentimiento y finalidad. Críticas y propuestas hacia una regulación de la protección de datos personales en Ecuador	Felipe Nicolás Roldán Carrillo [17]	Protección de datos personales
E8	Elaboración de la guía de implantación de las normas prioritarias del esquema gubernamental de seguridad de la información EGSI en las entidades de la administración pública central	Cáceres Tarco, Carolina Elizabeth, & Mena González, Cristina Elizabeth [25]	Protección de datos personales

Todos los estudios primarios tenían un enfoque o tema en relación con los mecanismos para la Recolección y Procesamiento Seguro de PII. El enfoque de cada documento también se registró en la Tabla 1.

El enfoque de cada artículo se agrupó en categorías más amplias para permitir una clasificación simplificada de los temas de los estudios primarios.

Los estudios que se centraron en las leyes de protección de datos se agruparon em la categoría ley de protección. Los estudios que tenían un enfoque relacionado con la Recolección y Procesamiento Seguro de Información de Identificación Personal (PII) se agruparon en la categoría de recolección y procesamiento de datos.

2. METODOLOGÍA

En esta sección se identificarán las metodologías utilizadas para la implementación y desarrollo del prototipo experimental, así como también las herramientas que darán paso a la solución de la problemática planteada.

2.1 Selección de Herramientas

Este trabajo seguirá el enfoque del paradigma de la investigación en ciencias del diseño Design Science Research (DSR por sus siglas en inglés) [18], el mismo que tiene un enfoque en el paradigma de la resolución de problemas. Otra metodología utilizada fue la metodología SCRUM que complementa al DSR en la parte del prototipo. Esta metodología permitirá una secuencia de pasos para ir construyendo el prototipo final del proyecto.

Las herramientas que serán utilizadas para la implementación del componente se describen en las siguientes subsecciones.

2.1.1 Design Science Research (DSR)

La DSR procura mejorar la comprensión humana mediante la creación de artefactos innovadores y la generación del diseño del conocimiento mediante soluciones innovadoras a problemas del mundo real representados por constructos, modelos, métodos e instancias [18]. El objetivo de la DSR es generar conocimiento de cómo pueden y deben construirse o disponerse las cosas, por medio del diseño, a través de la acción humana y que permitan lograr el cumplimiento de un conjunto de objetivos deseados.

El entorno puede definirse como el espacio del problema en el que residen los fenómenos de interés. Está compuesto por personas, organizaciones y tecnologías. En él se hallan los objetivos, las tareas, los problemas y las oportunidades que definen las necesidades de las partes interesadas de la organización. Estas necesidades se estiman y evalúan en el contexto de las estrategias organizativas, la estructura, la cultura y los procesos de trabajo existentes. Se instalan en correspondencia con la infraestructura tecnológica, aplicaciones, las arquitecturas y las capacidades de desarrollo. Todo lo mencionado finalmente define el problema de investigación percibido por el investigador. El siguiente cuadro de la Figura 6 describe un framework a seguir dentro del DSR:

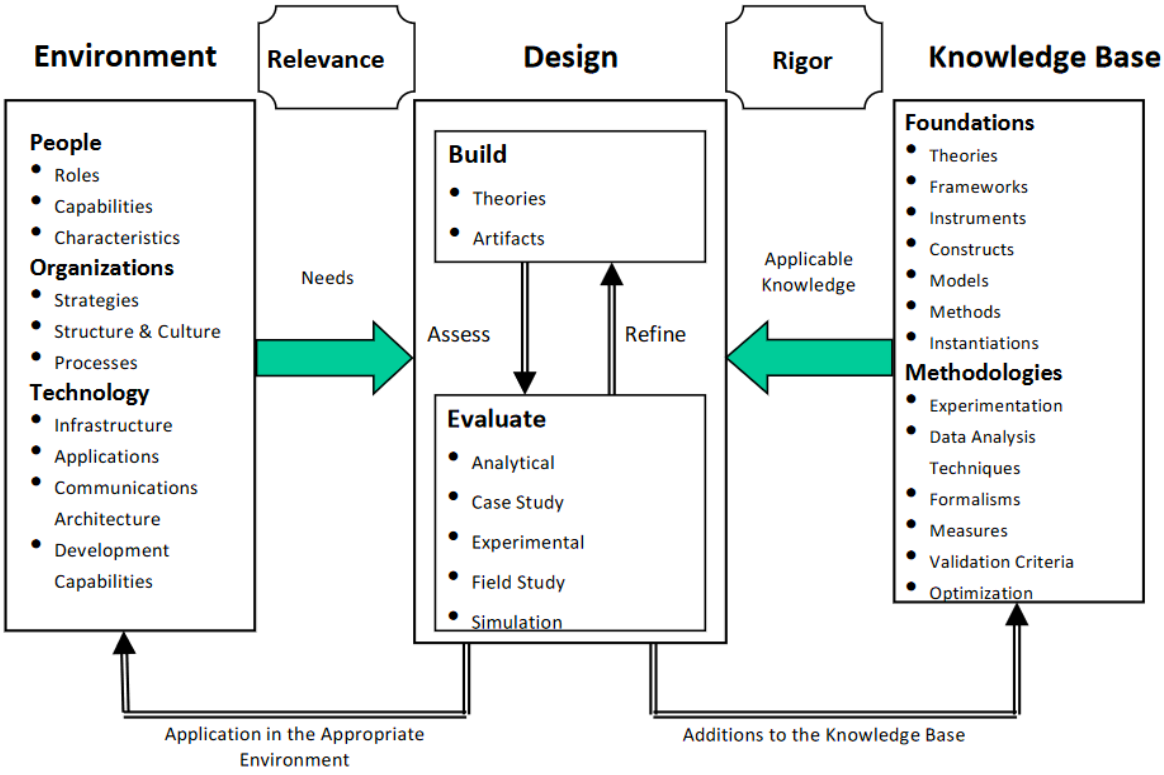

Figura 6: DSR framework [18].

Tras el análisis correspondiente de la DSR y su framework, el actual trabajo perseguirá las siguientes actividades para dar cumplimiento de los objetivos y alcance planteados en las secciones anteriores. Estas actividades se describen a continuación:

- **Actividad 1: Identificación de los problemas existentes y motivación.** Los problemas identificados en la presente investigación radican en la escasez de soluciones existentes durante la recolección y procesamiento de los datos personales de las empresas u organizaciones pertenecientes a algún Estado. A su vez, se presenta otro problema al no tener una referencia clara respecto a los requerimientos funcionales y no funcionales vinculados con la seguridad y privacidad de los datos personales de los individuos durante su recolección y procesamiento. Finalmente, no existen prototipos de sistemas o arquitecturas apoyadas en las leyes de la LOPDP, RGPD y EGSI, que guíen al responsable del Tratamiento junto con su Equipo Técnico durante la recolección y procesamiento de los datos personales. Por lo que es primordial implementar un prototipo inspirado en las leyes ya mencionadas con sus respectivos requerimientos funcionales y no funcionales que faciliten la comprensión y ejecución del tratamiento de los datos personales.
- **Actividad 2: Definir los objetivos de una solución.** Dentro de las posibles soluciones, se examina la implementación de mecanismos de control de acceso

durante la recolección y procesamiento de los datos personales de los individuos, garantizando su seguridad y privacidad. Los mecanismos implementados serán prototipos experimentales.

- **Actividad 3: Diseño y desarrollo.** Una vez diseñado el prototipo experimental, este será ejecutado en entornos controlados sin que esto implique el despliegue de estos en alguna infraestructura productiva, ya sea pública o privada.
- **Actividad 4: Demostración.** Esta actividad permitirá, tras la ejecución del prototipo experimental, resolver los casos o problemas presentes durante la recolección y tratamiento de los datos personales. Este paso se lo realizará a través de diferentes procesos como: experimentación, simulación, estudio de caso, pruebas y demás actividades relacionadas.
- **Actividad 5: Evaluación.** La evaluación permitirá medir lo bien que el prototipo experimental apoya a las soluciones de los problemas. Dentro de esta actividad se compararán los objetivos de las soluciones a los problemas planteados con los resultados reales observados de la ejecución del prototipo experimental durante la recolección y tratamiento de los datos personales. Una vez finalizada esta actividad, él o los investigadores podrán decidir si realizan un paso repetitivo en la actividad 3 para intentar mejorar la eficacia del prototipo o continuar con la comunicación y dejar las mejoras para proyectos posteriores.
- **Actividad 6: Comunicación.** La finalidad de esta actividad es comunicar todos los aspectos del problema y del prototipo experimental diseñado junto con los resultados obtenidos.

Las actividades planteadas están basadas en el modelo de proceso de la metodología DSR [26] y cuyo modelo se observa en la siguiente Figura 7:

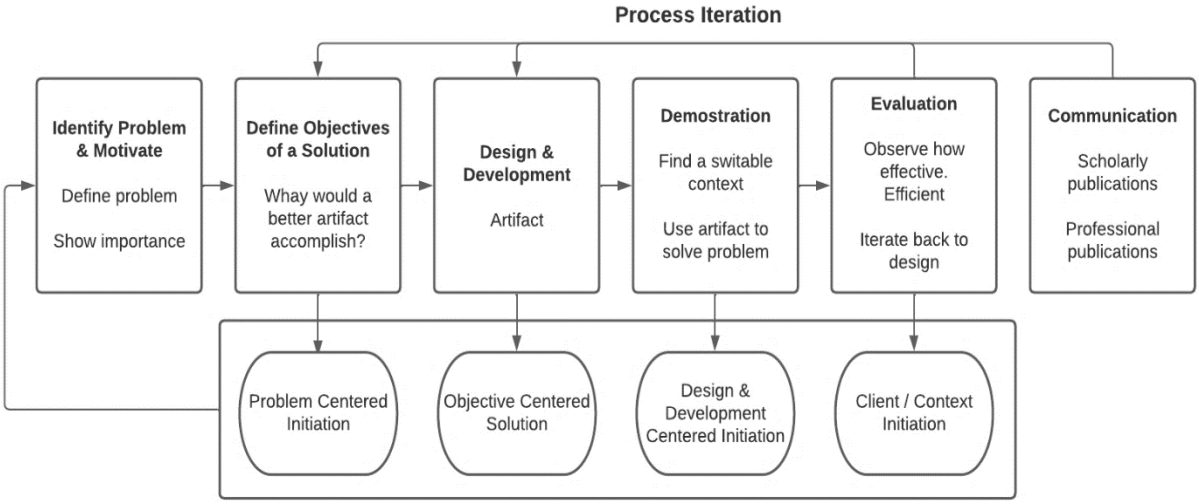

Figura 7: Modelo de proceso de la metodología DSR [26].

El tipo de trabajo actual está basado en teorías prescriptivas (también conocidas como teorías del diseño) que proporcionan instrucciones sobre "cómo hacer algo". Esto permite analizar y describir el proceso que se llevará a cabo durante la construcción del prototipo experimental, haciendo visible su modelización.

2.1.2 Metodología SCRUM

SCRUM es una metodología ágil utilizado para el desarrollo de software que enfatiza la colaboración, la flexibilidad y el desarrollo iterativo. Se basa en los principios de transparencia, fiscalización y adecuación. El proceso de Scrum se divide en ciclos de trabajo llamados sprints, que duran entre 1 y 4 semanas. Cada sprint comienza con una reunión de planificación de sprint, en la que el equipo de desarrollo se reúne con el propietario del producto para definir el trabajo que se realizará durante el sprint.

2.1.3 Microsoft Threat Modeling Tool

Actualmente, está altamente aceptado que las buenas prácticas de seguridad de software estén integradas en todo el ciclo del desarrollo de un programa. Una de las mejores formas para la seguridad de software es el modelamiento de amenazas (threat modeling). Es necesario utilizar una herramienta de modelamiento de amenazas durante la etapa de diseño del programa debido a que ayuda a reducir los errores de forma significativa antes de que la aplicación del programa sea implementada.

Microsoft Threat Modeling Tool es el elemento central del Microsoft Security Development Lifecycle (SDL). Este programa permite identificar y mitigar problemas potenciales de seguridad en una etapa temprana, cuando estos tienen una resolución costo-efectiva baja. Esta herramienta puede ser utilizada por todos los desarrolladores, sin tener que ser

expertos en seguridad; ya que, posee una guía clara en crear y analizar modelos de amenaza [27].

2.1.4 PYTHON

Python es un lenguaje de programación de alto nivel interpretado y orientado a objetos. Es un lenguaje simple, legible y fácil de aprender. Puede ser utilizado para el desarrollo web utilizando diferentes frameworks. Además, es flexible lo que permite a los desarrolladores la creación de aplicaciones web. Python es utilizado en una amplia variedad de aplicaciones, desde el desarrollo web hasta el análisis de datos, la inteligencia artificial y el aprendizaje automático. Es un lenguaje de programación de propósito general que se puede utilizar en muchos campos diferentes

2.1.5 Flask

Flask es un framework que puede ser fácilmente extendido, disponible para una gran cantidad de usos que facilitan la implementación de funcionalidades. Para la recolección de información es necesaria la creación de un formulario. Flask utiliza el motor de plantillas Jinja 2 para generar un HTML dinámico a partir de dichas plantillas. Para enviar la información es necesario configurar el framework para que pueda recibir las solicitudes HTTP del formulario web. La ruta debe especificar el método HTTP que se va a utilizar para enviar los datos del formulario (normalmente este es POST), así como el nombre de la función que se encarga de procesar lo datos.

2.1.6 MongoDB

MongoDB es una base de datos no relacional que no utiliza tablas, las bases de datos NoSQL emplean modelos de datos que pueden ser diferentes dependiendo del proyecto que se vaya a aplicar. Por otro lado, esta base de datos es escalable lo que significa que se puede añadir flexibilidad en esquema de datos dinámicos y permite trabajar con datos estructurados y no estructurados. Una de las principales ventajas es el alto rendimiento en ambientes con gran cantidad de datos, tráfico de lecturas y escritura.

2.2 Descripción de la Solución

Para el diseño de la solución planteada, se propone las siguientes subsecciones que se mencionan como se implementarán las herramientas que faciliten el trabajo de la recolección de los datos personales.

El diagrama de componentes representa la parte modular del subsistema a desarrollar para la recolección y procesamiento de PII. El diagrama de componentes se define en la siguiente Figura 8.

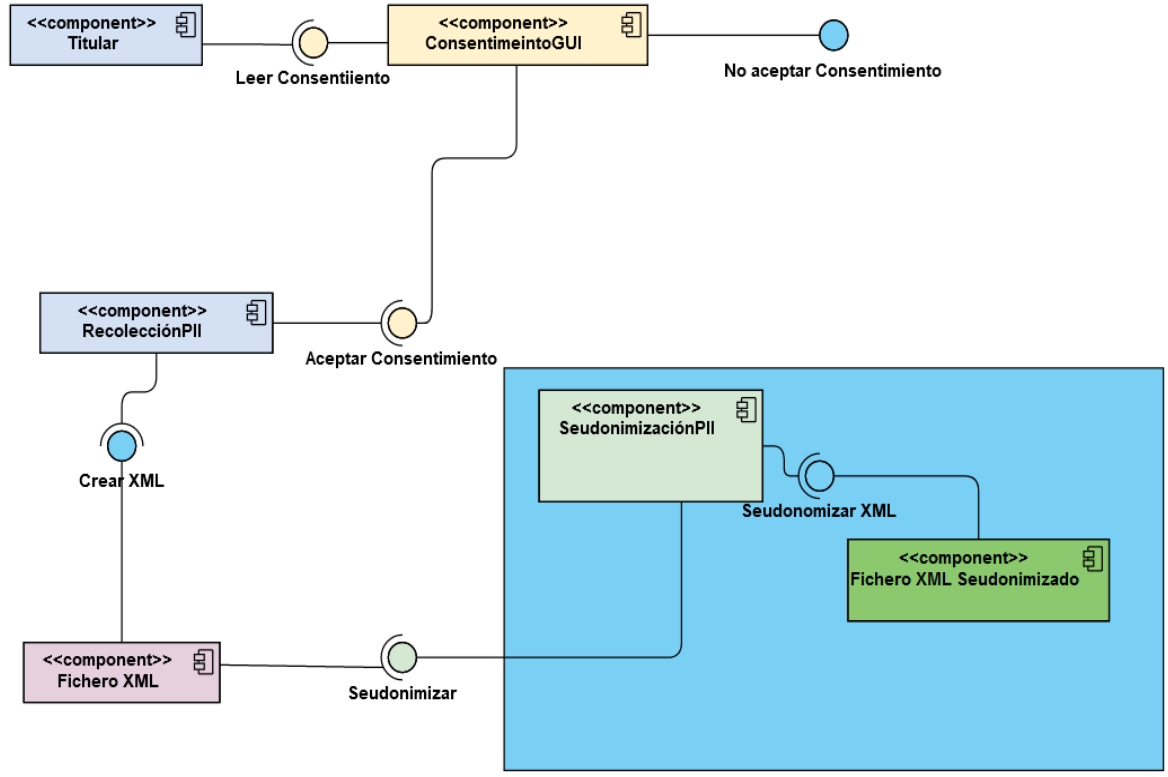

Figura 8: Diagrama de componentes para la recolección y procesamiento de PII.

[Autor: William Suntaxi]

2.2.1 Descripción de diagrama de componentes

Como se observa en la Figura 8, el diagrama de componentes para la recolección y procesamiento de PII, es necesario describir, cada uno de los componentes para su posterior implementación. La Tabla 2 muestra la información de cada uno de los componentes.

Tabla 2. Descripción de diagrama de componentes.

[Autor: William Suntaxi]

Componente	Descripción
Titular	Ilustra la primera pieza del proceso de recolección de los datos; es el componente que posiblemente entregaría sus datos personales.
ConsentimientoGUI	Formulario encargado de mostrar y pedir la validación del titular junto con el consentimiento de entrega de datos.

RegistroGUI	Formulario encargado de recopilar la información de la organización y de la aceptación del componente de consentimiento.
Minimización PII	Componente en el cual se realiza la minimización de la información de los datos recolectados del Titular.
SeudominizaciónPII	Este componente se encarga de manejar la pseudominización para proteger la información privada de los usuarios. Este componente se encarga de reemplazar los datos personales originales con valores ficticios.
CifradoClaveSecretaPI	Componente en el cual se aplica la técnica de seudonimización de cifrado de clave secreta.
Recolección PII	Este componente se encarga de la recolección de datos personales, además se encargará del cifrado de la información, control de acceso.

2.2.2 Pseudocódigo de diseño de solución para recolección y procesamiento de PII

En esta subsección se describe el pseudocódigo con los pasos del proceso de recolección y procesamiento de PII. El pseudocódigo se muestra a continuación:

```
inicio
    1.  función principal():
    2.      si request.method == 'POST':
    3.          # Obtenemos la cédula del formulario
    4.          cedula = request.form['ci']
    5.          # Verificamos si la cédula es válida
    6.          si verificar_cedula(cedula):
    7.              data = request.form
    8.              session["datos"] = data
    9.              return redirect(url_for("aceptarTerminos"))
    10.         sino:
    11.             # Si no es válida, volvemos a mostrar el formulario
    12.             return render_template('form.html', error=True)
    13.     sino:
    14.         # Si es una petición GET, simplemente mostramos el formulario
    15.         return render_template('form.html')
fin
```

```
inicio
    1.  función aceptarTerminos()
    2.  si request.method == 'POST' entonces
    3.          return render_template('notificación.html')
    4.  sino si request.method == 'GET' entonces
    5.          return redirect(url_for('formulario'))
    6.  fin si
```

```
    7.  fin función

    8.  función solicitar_permiso()
    9.  si request.method == 'POST' entonces
    10.     datos = request.form
    11.     # procesar los datos y enviar una respuesta
    12. fin si
    13. fin función

    14. # definir las rutas
    15. app.route('/terminos_codiciones', methods=['POST', 'GET']) -> aceptarTerminos
    16. app.route('/solicitar_permiso', methods=['POST']) -> solicitar_permiso
fin
```

2.2.3 Solución para la recolección de los datos personales

Para la recolección de datos personales se implementará un aplicativo web cuyo diseño deberá cumplir con las siguientes funcionalidades:

- La aplicación mostrará un contrato en formato digital, el cual contendrá las bases legales y cómo será el tratamiento de los datos personales del titular.
- Una vez se haya aceptado el contrato, la aplicación desplegará una ventana en la cual se tomarán los datos personales. Para el desarrollo de esta funcionalidad se debe tener en cuenta los principios de pertenencia y minimización.
- Luego de la recolección de datos, se mostrará una ventana de validación de estos datos, la cual servirá para detectar algún error durante el ingreso de datos.
- Finalmente, se muestra un mensaje de almacenamiento correcto de los datos y el proceso finalizará.

Este proceso se verá respaldado por el responsable del tratamiento, quien es el representante de la organización.

2.2.4 Requerimientos para la recolección y procesamiento de datos personales.

Tabla 3. Requerimientos del sistema

[Autor: William Suntaxi]

ID.	Requerimiento	Tipo de Requerimiento

REQ_001	El sistema se encargará de controlar el acceso y dará paso únicamente a usuarios debidamente autorizados. Los usuarios ingresarán al sistema por medio de un nombre de usuario y contraseña.	Funcional
REQ_002	Se enviará un mensaje de alerta al encargado de la administración del sistema cuando se provoque algunos de los siguientes acontecimientos: - Registro de una nueva cuenta de usuario - Inicio de sesión - Dos o más intentos fallidos al momento de ingresar las credenciales - Cambio de contraseña	Funcional
REQ_003	Posterior al inicio de sesión, el sistema mostrará una ventana en la cual se muestre el consentimiento o contrato escrito.	Funcional
REQ_004	La ventana que muestre el consentimiento tendrá dos opciones en la cual se acepta o no dicho acuerdo.	Funcional
REQ_005	Tras la aceptación del acuerdo, el sistema mostrará una ventana con campos de registro para ingresar los datos personales de la persona.	Funcional
REQ_006	Los datos personales deberán ser relevantes para el propósito de su uso y, en la medida de lo necesario para dicho propósito, exactos, completos y actuales.	Funcional
REQ_007	El propósito de la recolección de datos se deberá especificar al momento que se realiza dicha recolección. Además, el uso de los datos se verá estrictamente limitado con el cumplimiento de los objetivos propuestos u otros que no tengan relación con el propósito original.	Funcional
REQ_008	El sistema deberá garantizar la seguridad de los datos, de manera que se preserve la integridad de estos y se impida el acceso o uso no autorizado de los datos durante su recolección o procesamiento. Si una persona	Funcional

	interviene en alguna de las fases del tratamiento de los datos personales, deberá cumplir con un acuerdo de confidencialidad con tiempo indefinido.	
REQ_009	Los datos deben ser exactos y, en caso de errores, deben ser corregidos y actualizados. Para ello se deberán implementar módulos del sistema correctoras necesarias y razonables que permitan modificar los datos inexactos o incompletos, para así poder garantizar la veracidad y seguridad de la información objeto de tratamiento.	Funcional
REQ_010	Si los datos se ingresaron correctamente, el sistema mostrará un mensaje de éxito.	Funcional
REQ_011	El sistema ejecutará de forma automática el proceso de minimización de los datos usando el patrón de diseño de pseudonimización.	Funcional
REQ_012	El sistema incluirá un proceso de pseudonimización por medio de la aplicación de clave secreta	Funcional
REQ_013	El sistema mostrará un mensaje de alerta al administrador del sistema cuando se hayan minimizado los datos ingresados sin error alguno durante el respectivo proceso.	Funcional
REQ_014	El proceso de minimización de los datos ingresados no deberá tardar más allá de 5 segundos	No Funcional
REQ_015	Durante la recolección y procesamiento, se deberá implementar dentro del sistema el protocolo HTTPS que permita establecer una conexión segura entre el servidor y el responsable, para que esta no pueda ser interceptada por personas no autorizadas	No Funcional

2.2.5 Propuesta de ingreso y procesamiento de datos

Una vez que el Titular ha proporcionado su consentimiento, es necesario describir en que formato se tomará la información que fueron proporcionados para su posterior tratamiento.

En la Figura 9 se explica el proceso de recolección de información desde que el Titular proporciona su información, una vez aprobado esta información se realiza una persistencia temporal utilizando un fichero XML para la transferencia de información y el ocultamiento de información confidencial utilizando técnicas de pseudonimización y posteriormente se realizará la persistencia temporal en una base de datos No Relacional.

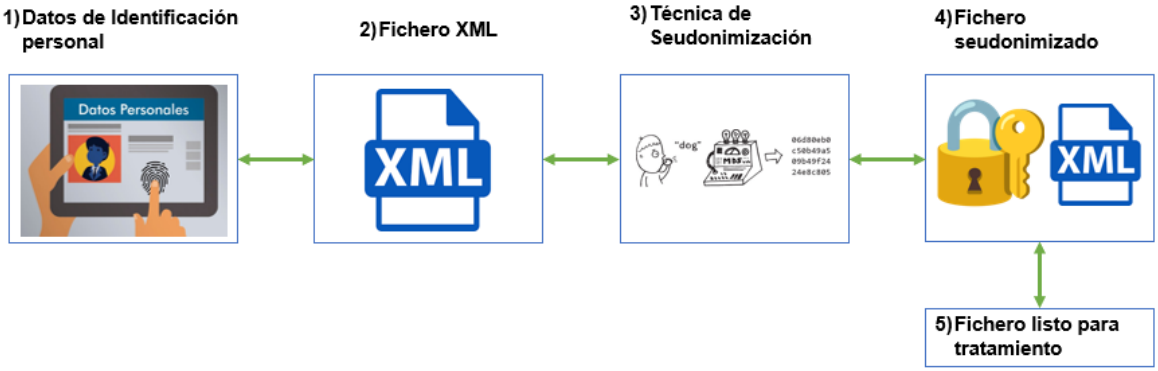

Figura 9: Proceso de recopilación y pseudonimización de datos de identificación personal. [Autor: William Suntaxi]

En cuanto a la estructura del archivo XML podría ser de la siguiente manera, tal como se muestra en la Figura 10.

```
<root>
<item type="dict">
<cedula type="str">1726487893</cedula>
<nombres type="str">CRISTHIAN FERNANDO</nombres>
<apellidos type="str">TOHASA PASPUEZAN</apellidos>
<sexo type="str">HOMBRE</sexo>
<condicion_ciudadano type="str">CIUDADANO</condicion_ciudadano>
<fecha_nacimiento type="str">1997-11-30</fecha_nacimiento>
<lugar_nacimiento type="str">PICHINCHA/QUITO/SAN ROQUE</lugar_nacimiento>
<nacionalidad type="str">ECUATORIANA</nacionalidad>
<estado_civil type="str">SOLTERO</estado_civil>
<codigo_dactilar type="str">E3343V2242</codigo_dactilar>
<domicilio type="str">PICHINCHA/QUITO/CHILLOGALLO</domicilio>
<calles_domicilio type="str">CALE E CALLE A BUENA VENTURA</calles_domicilio>
<numero_casa type="str">MZ 85</numero_casa>
<telefono type="str">0987149347</telefono>
<correo_electronico type="str">cristhian.tohasa@epn.edu.ec</correo_electronico>
<estado_afiliado type="str">ACTIVO</estado_afiliado>
<ruc_patronal type="str">1792787289001</ruc_patronal>
<sector type="str">OTROS</sector>
<razon_social type="str">CONSORCIO ALCANTARILLADOS</razon_social>
<origen type="str">HL</origen>
<periodo_desde type="str">2018-09-01</periodo_desde>
<periodo_hasta type="str">2018-10-01</periodo_hasta>
<imposiciones type="int">0</imposiciones>
<dias type="int">16</dias>
</item>
```

Figura 10: Estructura XML para la transferencia de datos.

[Autor: William Suntaxi]

2.2.6 Mapeo de requerimientos entre leyes y la literatura

Una vez planteados los requerimientos que el sistema debe tener, es importante mapear estos mecanismos o principios junto con las leyes y las investigaciones encontradas en la revisión sistemática de la literatura.

La Tabla 4 muestra la información de todo el mapeo y la información antes mencionada.

Tabla 4. Mapeo entre requerimientos, leyes y literatura.

[Autor: William Suntaxi]

Req. No.	**Mecanismo**	**Instrumentos**		
		GDRP	LOPDP	EGSI
REQ_001	Licitud, lealtad y transparencia	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	Secciones: 1.1, 4.1, 4.2, 5.1, 5.2, 5.3, 5.4.
REQ_002	Licitud, lealtad y transparencia	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41,	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49,	Secciones: 1.1, 4.1, 4.2, 5.1, 5.2, 5.3, 5.4.

		42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	
REQ_003	Limitación de la finalidad, Licitud del tratamiento	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	Secciones: 1.1, 2.1, 3.2, 4.1, 4.2, 4.3, 5.1, 5.2, 5.3, 5.4, 8.1, 8.2, 8.3, 9.2, 10.1, 10.2.
REQ_004	Limitación de la finalidad, Licitud del tratamiento	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	Secciones: 1.1, 2.1, 3.2, 4.1, 4.2, 4.3, 5.1, 5.2, 5.3, 5.4, 8.1, 8.2, 8.3, 9.2, 10.1, 10.2.

REQ_005	Licitud del tratamiento	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	Secciones: 1.1, 4.1, 4.2, 5.1, 5.2, 5.3, 5.4.
REQ_006	Minimización	Artículos: 5, 25, 47, 89	Artículos: 7, 9, 10, 21, 26	Secciones: 4.1, 4.2, 4.3, 9.2, 10.3, 11.1.
REQ_007	Licitud del tratamiento	Artículos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42, 44, 47, 51, 55, 56, 57, 58, 60, 62, 64, 71, 77, 79, 80, 81, 82, 85, 86, 87, 88, 89, 91, 94, 95, 98.	Artículos: 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 53, 56, 57, 58, 65, 67, 68, 69, 70, 76.	Secciones: 1.1, 4.1, 4.2, 5.1, 5.2, 5.3, 5.4.
REQ_008	Integridad y Confidencialidad	Artículos: 5, 28, 32, 38, 76.	Artículos: 10, 14, 18, 19, 27, 29, 37, 70	Secciones: 4.2, 4.3, 5.2, 5.3, 5.4, 7.2, 8.1, 8.2,

				8.3, 9.2, 14.2.
REQ_009	Exactitud	Artículos: 28, 35, 41, 46, 47, 51, 57, 60, 63, 66, 70, 74, 78, 85, 97.	Artículos: 10, 19, 29	Secciones: 1.1, 4.2, 5.1, 5.3, 8.3, 9.2, 10.1, 13.1.
REQ_010	Exactitud	Artículos: 28, 35, 41, 46, 47, 51, 57, 60, 63, 66, 70, 74, 78, 85, 97.	Artículos: 10, 19, 29	Secciones: 1.1, 4.2, 5.1, 5.3, 8.3, 9.2, 10.1, 13.1.
REQ_011	Minimización	Artículos: 5, 25, 47, 89	Artículos: 7, 9, 10, 21, 26	Secciones: 4.1, 4.2, 4.3, 9.2, 10.3, 11.1.
REQ_012	Encriptación	Artículos: 6, 32, 34.	Artículos: 37	Secciones: 1.1, 5.2, 5.4, 6.1, 7.2, 9.1, 10.1, 14.2.
REQ_013	Minimización	Artículos: 5, 25, 47, 89	Artículos: 7, 9, 10, 21, 26	Secciones: 4.1, 4.2, 4.3, 9.2, 10.3, 11.1.
REQ_014	Minimización	Artículos: 5, 25, 47, 89	Artículos: 7, 9, 10, 21, 26	Secciones: 4.1, 4.2, 4.3, 9.2, 10.3, 11.1.
REQ_015	Confidencialidad	Artículos: 5, 28, 32, 38, 76.	Artículos: 10, 30, 31, 44, 45, 47, 70.	Secciones: 4.2, 4.3, 5.2, 5.3, 5.4, 7.2, 8.1, 8.2, 8.3, 9.2, 14.2.

2.2.7 Propuesta de recolección y transferencia de datos personales

Los diagramas de secuencia modelan las interacciones entre los objetos en un solo caso de uso. Como se muestra en la Figura 11, el proceso de recolección de información inicia el usuario, dando información mediante un formulario para el cual se realiza la confirmación email mediante un OTP para la autorización, una vez se ha realizado la autorización se realiza la persistencia en un archivo XML. Este archivo contendrá la información recolectada del formulario. Una vez se tenga el archivo XML se identifica la información

privada proporcionada por el cliente para proceder a pseudonimizarla, utilizando un método de hash con un salting adicional.

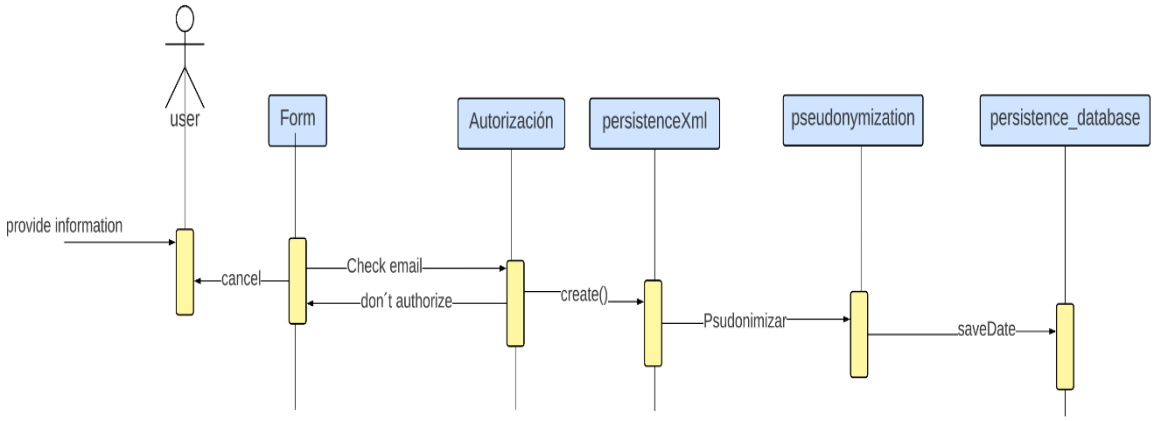

Figura 11: Diagrama de secuencias para el Componente 1.
[Autor: William Suntaxi]

El diagrama clases es una herramienta de un modelado que se utiliza en el desarrollo de software para la visualización de la estructura de un sistema orientado a objetos, lo que facilita la visualización y la escalabilidad del software. Además, permite entender de manera clara el flujo de información, como se muestra en la Figura 12.

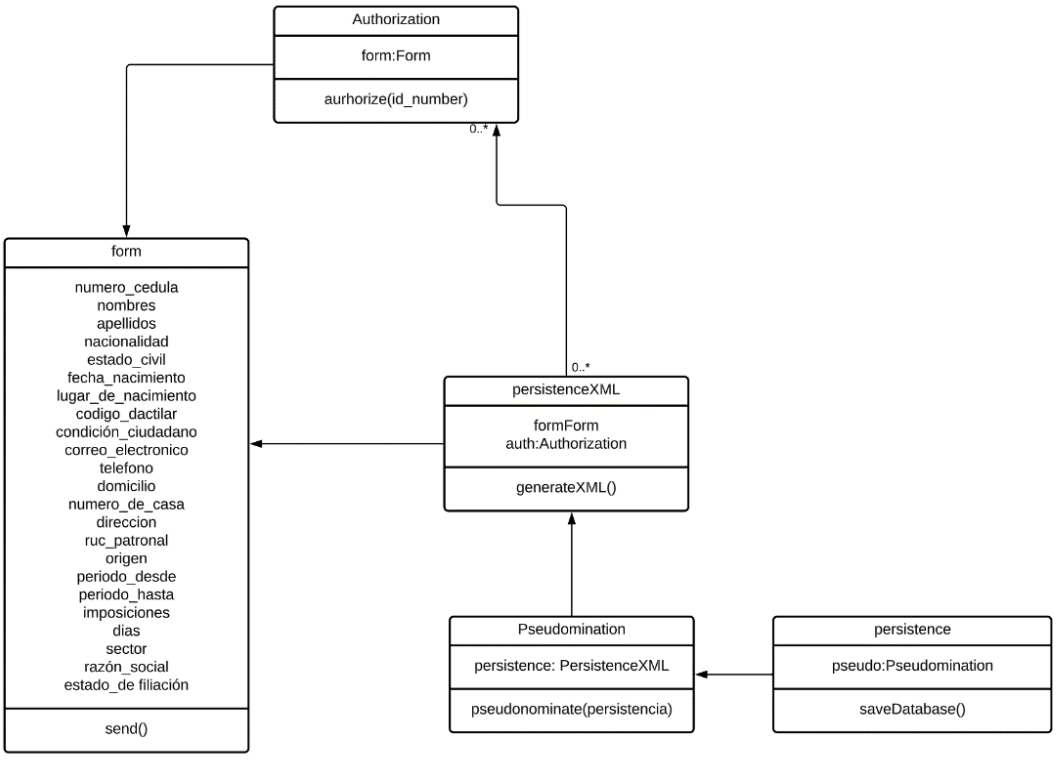

Figura 12: Diagrama de clases para el Componente 1.
[Autor: William Suntaxi]

En la Figura 13 se muestra el diagrama de actividades para el proceso de recolección de información, el proceso inicia proporcionando la información del usuario a un formulario renderizado por flask, una vez es entregado se realiza la validación por parte del sistema, en caso de que la información proporcionada por el cliente no se valide se vuelve a solicitar la información, caso contrario muestra las políticas de privacidad si el usuario acepta dichas políticas. En caso de negativa se vuelve a solicitar la información, caso contrario se genera un XML con los datos personales del usuario, una vez generado este archivo se procede a la pseudonimización de la información, para posteriormente minimizar la información.

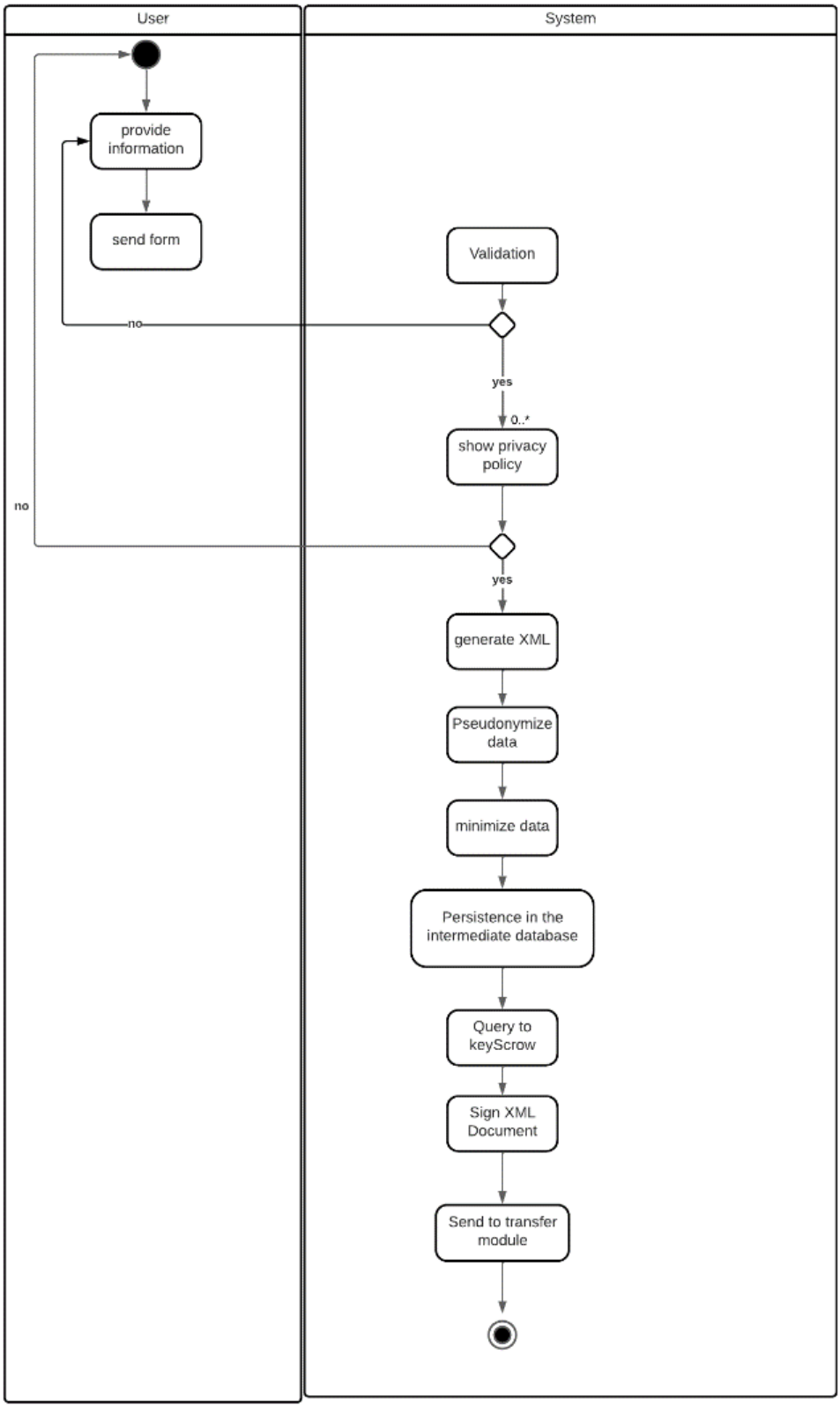

Figura 13: Diagrama de actividades para el proceso de recolección de información.

[Autor: William Suntaxi]

En la Figura 14 se describe la consulta por una entidad externa. El proceso inicia cuando la entidad externa accede a la información, para lo cual se debe realizar un login utilizando certificados digitales, una vez validados los certificados digitales utilizando un keyscrow se valida la autenticidad y se verifican las firmas digitales del archivo XML para mostrar la información de la base de datos.

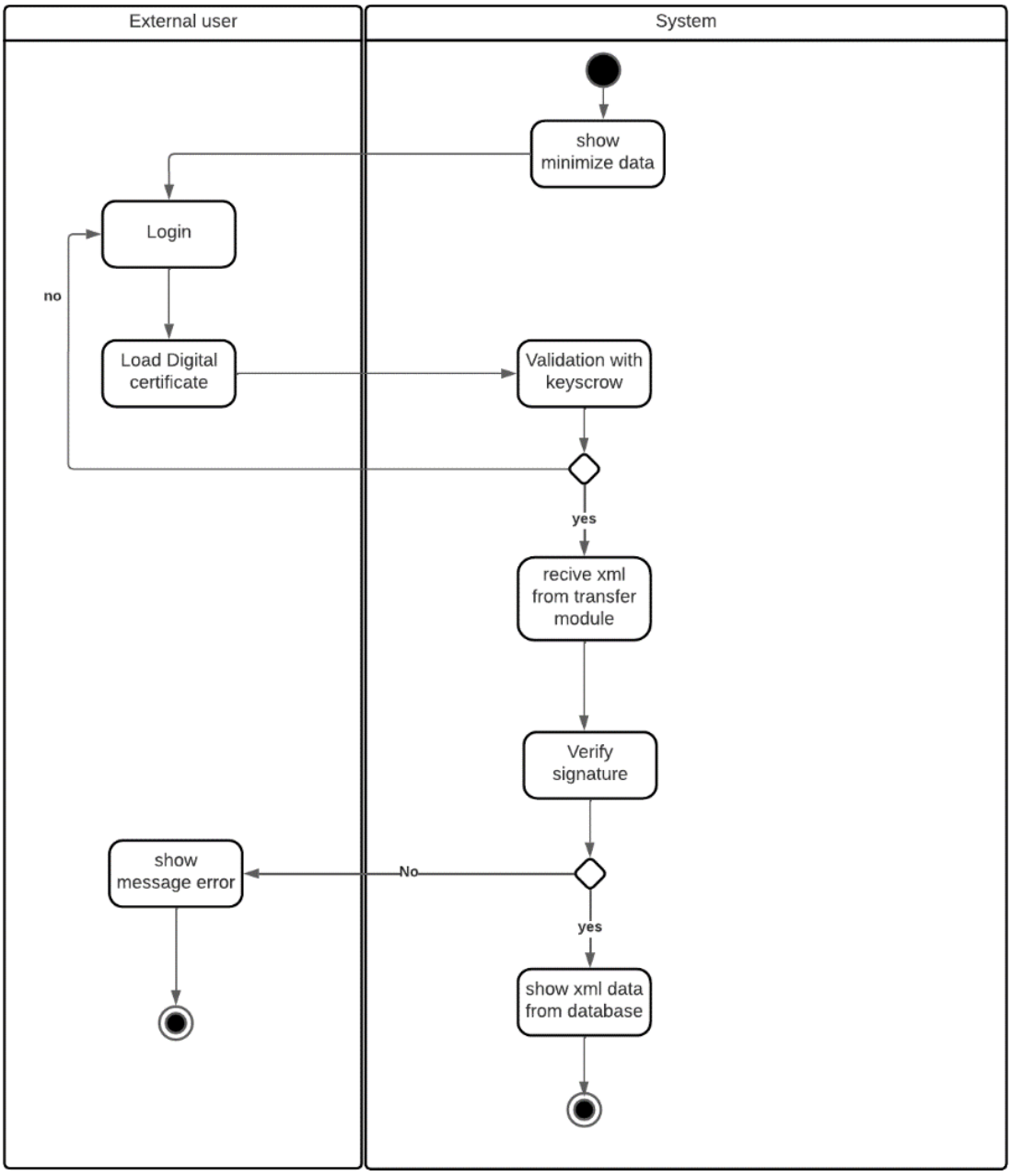

Figura 14: Diagrama de actividades para el proceso de consulta por una entidad externa. [Autor: William Suntaxi]

En la Figura 15 se muestra la arquitectura del módulo de recolección de información el cual inicia con un factor humano proporcionando información para luego ser trasferida a una base de datos intermedia que contiene información pseudonimizada y que estará disponible para consultas de entidades externas. Finalmente, la entidad externa debe realizar una autorización previa utilizando certificados digitales para el acceso a la base de datos.

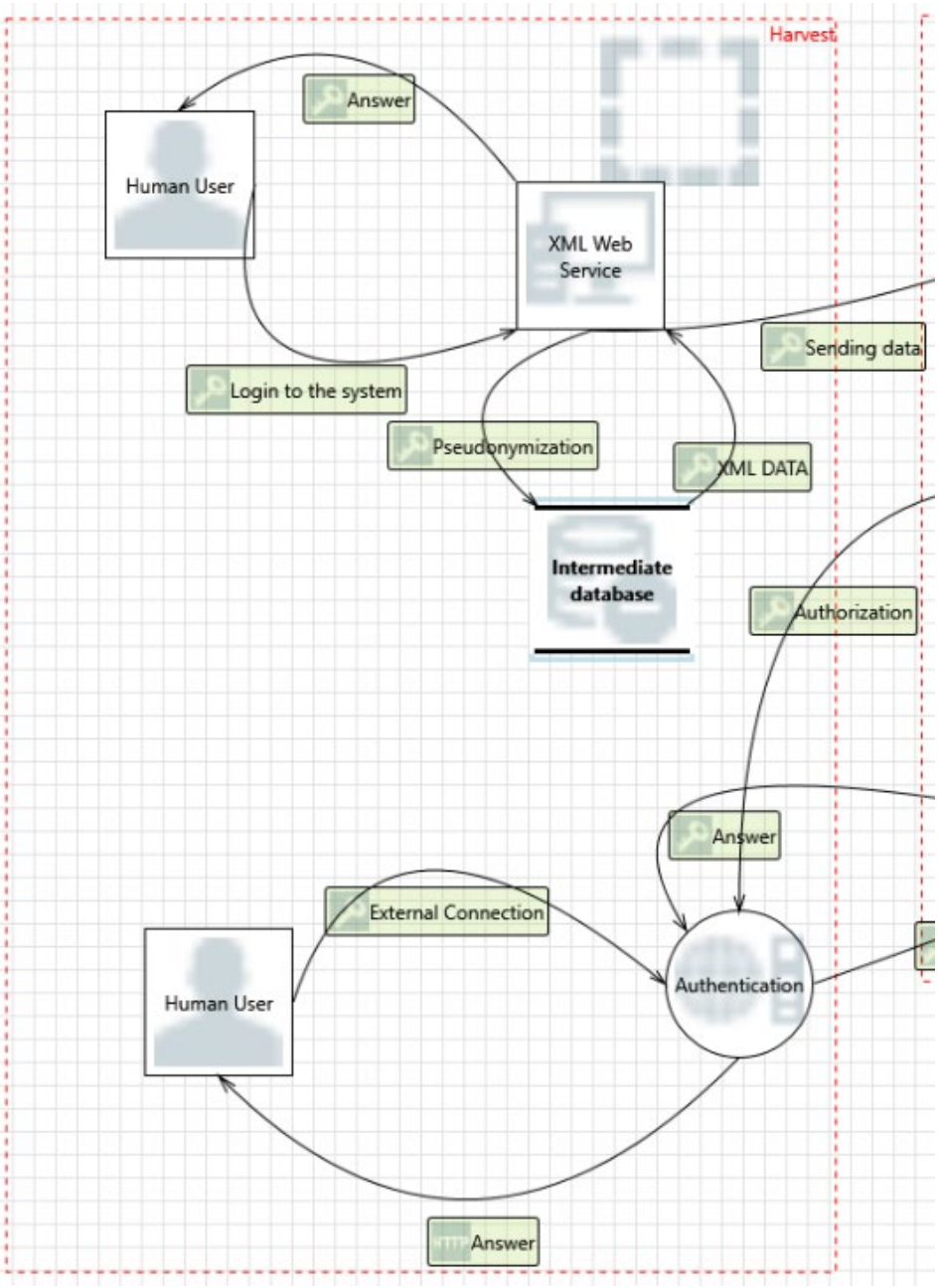

Figura 15: Diagrama Arquitectura del Componente de Recolección y Transferencia de Información del Proyecto TIC.

[Autor: William Suntaxi]

3. EVALUACIÓN, CONCLUSIONES Y RECOMENDACIONES

3.1 Pruebas

Las pruebas que se realizarán en la investigación para obtener los resultados de desempeño, efectividad y eficiencia fueron la siguientes:

- Pruebas funcionales: son utilizadas en el desarrollo de software para comprobar que el sistema funciona según lo provisto y que cumple con los requerimientos planteados.
- Pruebas de rendimiento: este tipo de pruebas para evaluar cómo el sistema responde ante diferentes cargas y situaciones de uso.

A continuación, se muestran los resultados obtenidos con las pruebas antes mencionadas.

3.1.1 Pruebas Funcionales

Se registra la información de manera correcta en el archivo XML, junto con la información del formulario.

Algoritmo de validación para la cédula de ciudadanía:

```
@app.route('/formulario', methods=['GET', 'POST'])
def hello_world():
    if request.method == 'POST':


        if verificar_cedula(cedula):
            data = request.form
            session["datos"] = data
            return redirect(url_for("aceptarTerminos"))
        else:

            return render_template('form.html', error=True)
    else:
        # Si es una petición GET, simplemente mostramos el formulario
        return render_template('form.html')
```

Figura 16: Algoritmo de validación para la cédula de ciudadanía.

[Autor: William Suntaxi]

Algoritmo de verificación para la cédula de ciudadanía:

```python
def existe_cedula(cedula):
    # Obtener todos los documentos de la colección que contienen el "salt" y el hash de la cédula
    cursor = mongo.db.coleccion_minimizados.find({}, {'salt': 1, 'id': 1})
    # Recorrer cada documento y obtener el "salt" y el hash almacenados
    for documento in cursor:
        salt = documento['salt']
        hash_almacenado = documento['id']
        # Concatenar el "salt" con la cédula que se está verificando
        cedula_salt = salt + cedula
        # Generar el hash de la cédula con la "salt" concatenada
        cedula_hash = hashlib.sha256(cedula_salt.encode()).hexdigest()
        # Comparar el hash generado con el hash almacenado en el documento
        if cedula_hash == hash_almacenado:
            return True
    return False
```

Figura 17: Algoritmo de verificación para la cédula de ciudadanía

[Autor: William Suntaxi]

En la Figura 18, se muestran operaciones CRUD para el acceso a la información:

```python
def insertarDatos(datos):
    mongo.db.coleccion_minimizados.insert_one(datos)

def mostrarDatos():
    # Realiza una consulta para obtener todos los documentos de la colección
    cursor = mongo.db.coleccion_minimizados.find()
    print(cursor)
    for documento in cursor:
        print(documento)

def consultaDatos():
    cursor = mongo.db.coleccion_minimizados.find()
    datos = []
    for documento in cursor:
        datos.append(documento)
    print(datos)
    return  datos
```

Figura 18: Operaciones CRUD para consulta de base de datos.

[Autor: William Suntaxi]

En la Figura 19, se muestra el registro de información:

```
<root>
<item type="dict">
<cedula type="str">1726487893</cedula>
<nombres type="str">CRISTHIAN FERNANDO</nombres>
<apellidos type="str">TOHASA PASPUEZAN</apellidos>
<sexo type="str">HOMBRE</sexo>
<condicion_ciudadano type="str">CIUDADANO</condicion_ciudadano>
<fecha_nacimiento type="str">1997-11-30</fecha_nacimiento>
<lugar_nacimiento type="str">PICHINCHA/QUITO/SAN ROQUE</lugar_nacimiento>
<nacionalidad type="str">ECUATORIANA</nacionalidad>
<estado_civil type="str">SOLTERO</estado_civil>
<codigo_dactilar type="str">E3343V2242</codigo_dactilar>
<domicilio type="str">PICHINCHA/QUITO/CHILLOGALLO</domicilio>
<calles_domicilio type="str">CALE E CALLE A BUENA VENTURA</calles_domicilio>
<numero_casa type="str">MZ 85</numero_casa>
<telefono type="str">0987149347</telefono>
<correo_electronico type="str">cristhian.tohasa@epn.edu.ec</correo_electronico>
<estado_afiliado type="str">ACTIVO</estado_afiliado>
<ruc_patronal type="str">1792787289001</ruc_patronal>
<sector type="str">OTROS</sector>
<razon_social type="str">CONSORCIO ALCANTARILLADOS</razon_social>
<origen type="str">HL</origen>
<periodo_desde type="str">2018-09-01</periodo_desde>
<periodo_hasta type="str">2018-10-01</periodo_hasta>
<imposiciones type="int">0</imposiciones>
<dias type="int">16</dias>
</item>
```

Figura 19: Registro de información.

[Autor: William Suntaxi]

En la Figura 20, se realiza la verificación del OTP registrado en la base de datos y se renderiza el login del aplicativo.

```python
@app.route("/verify_otp", methods=["GET", "POST"])
def verify_otp():
    if request.method == "POST":
        # Obtener el OTP ingresado por el usuario a partir del formulario
        otp = request.form.get("otp")
        usuario = session.get("usuario")
        users = mongo.db.login
        login_user = users.find_one({'username': usuario})
        # Verificar el OTP con la información almacenada en el servidor


        if otp == login_user['otp']:

            return redirect(url_for("obtener_datos"))
        else:
            # Si el OTP no es válido, mostrar un mensaje de error
            return render_template("verify_otp.html", error="OTP inválido")
```

Figura 2020: Verificación del OTP.

[Autor: William Suntaxi]

La Figura 21 muestra la estructura de un Login externo para el acceso a la información, se valida el usuario y la carga de los certificados digitales.

```
@app.route('/login_externa', methods= ['GET','POST'])
def login_externa():
    users = mongo.db.login
    #username = request.form['username']
    #login_user = users.find_one({'username': username})
    user_cert = request.files['file']
    cert = crypto.load_certificate(crypto.FILETYPE_PEM, user_cert.read())
    usuario = request.form['username']
    if usuario:
        login_user = users.find_one({'username': usuario})
        if login_user:
            subject = cert.get_subject()
            email = subject.emailAddress
            priv = login_user['privkey']
            priv_bytes = bytes(priv[2:-1], 'utf-8')
            pkey = crypto.load_privatekey(crypto.FILETYPE_PEM, priv_bytes)

            certUser = login_user['certificate']
            certUser_bytes = bytes(certUser[2:-1], 'utf-8')
            cert2 = crypto.load_certificate(crypto.FILETYPE_PEM, certUser_bytes)

            # Agregar el certificado autofirmado a la cadena de confianza
            store = crypto.X509Store()
            store.add_cert(cert)

            # Verificar el certificado con la cadena de confianza
            try:
                crypto.X509StoreContext(store, cert).verify_certificate()
                pkey.check()
                print('El certificado es válido.')
                # comparar huellas de los certificados
                if cert.digest("sha256") == cert2.digest("sha256"):
                    print("Los certificados son iguales")
                    otp = generate_otp()
                    send_otp_email(email, otp)
                    #Cuando se acepte la solicitud se debe enviar el nivel en una session
                    #se hace el update tomando en cuenta el nivel
                    users.update_one({'username': usuario}, {'$set': {'otp': otp}})
                    session["usuario"] = usuario
                    return  redirect(url_for('otp'))
                else:
                    print("Los certificados son diferentes")

            except crypto.Error as e:
                print('El certificado es inválido: {}'.format(str(e)))
            else:
                print("El usuario no existe")
        else:
            print("No se especificó un nombre de usuario")


    return  redirect(url_for('login'))
```

Figura 2121: Login externa.

[Autor: William Suntaxi]

En la Figura 22 se realiza la firma del archivo XML para el envío hacia el módulo de transferencia de información. Genera el XML de los datos recolectados y se procede a ser firmados para ser enviados al módulo de transferencia de información.

```python
@app.route('/confirmacion')
def confirmacion():
    datos = session.get("datos")
    root = ET.Element('root')
    for key, value in datos.items():
        element = ET.SubElement(root, key)
        element.text = value
    xml_str = ET.tostring(root, encoding='utf-8')
    with open('archivo.xml', 'wb') as f:
        f.write(xml_str)
    if bd_intermedia.existe_cedula(datos['ci']):
        return "Ya existe este usuario"
    else:
        insertarDatos(minimizar_datos(datos))
        #Obtener claves del keyscrow
        (pubkey, privkey) = obtenerClaves("william.suntaxi")
        xml_dict = {'xml_document': xml_str.decode('utf-8')}
        xml_firmado_william = firmar_xml(xml_dict,privkey,pubkey)
        cert = ('certificados/cert.pem', 'certificados/key.pem')
        #url = 'http://127.0.0.1:5000/recibirDanny'
        url = 'https://127.0.0.1/recibirDanny'
        headers = {'Content-type': 'application/json'}
        response = requests.post(url,
                                 data=json.dumps(xml_firmado_william),
                                 headers=headers,
                                 #cert=cert,
                                 verify=False
                                 )
        if response.status_code == 200:
            print("Datos enviados exitosamente.")
        else:
            print("Error al enviar los datos: %s" % response.text)

        return "La info se guardó"
    #return render_template('infoGuardada.html')
```

Figura 2222: Confirmación de envío de información.

[Autor: William Suntaxi]

En la Figura 23. Se muestra la información almacenada en una base de datos no relacional, como se puede observar se tiene los datos personales pseudonimizados, para posteriormente ser consultados por una entidad externa como se muestra en la Figura 23.

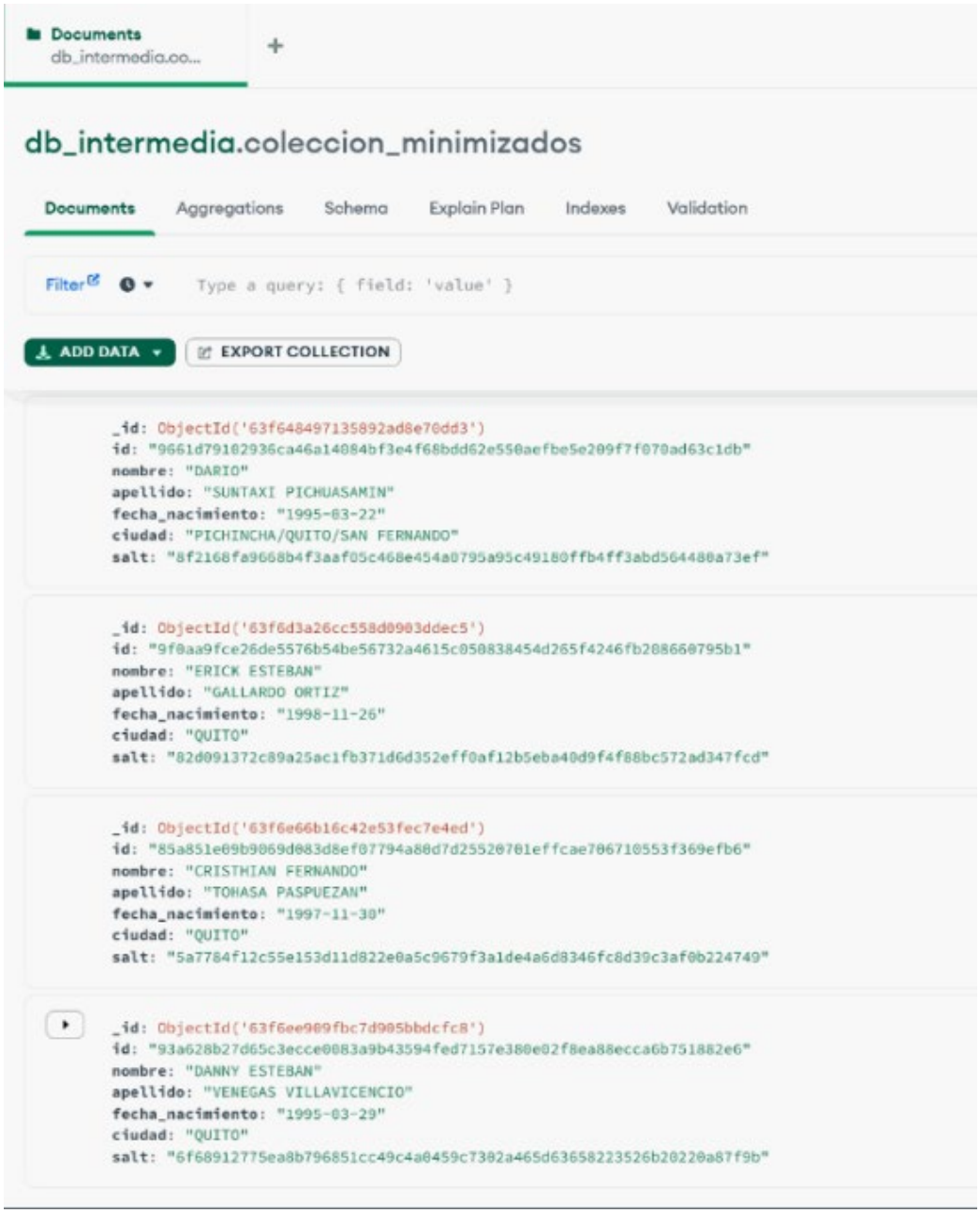

Figura 23: Registro de colección minimizada.

[Autor: William Suntaxi]

Codigo_id	Nombres	Apellidos	Fecha Nacimiento	Lugar de nacimiento
c3c9053e40952163c25dc1044510eaf291f0317acbd96453ae1accf72c7e029e	DARIO	SUNTAXI PICHUASAMIN	1988-05-24	Quito
e70a48bd821a02fa714cbf49ebf25433014745fe0f1a5d92eec1f29dec6006fd	Cristhian Fernando	Tohasa Paspuezan	1988-05-24	Quito
9661d79102936ca46a14084bf3e4f68bdd62e550aefbe5e209f7f070ad63c1db	DARIO	SUNTAXI PICHUASAMIN	1995-03-22	PICHINCHA/QUITO/SAN FERNANDO
9f0aa9fce26de5576b54be56732a4615c050838454d265f4246fb208660795b1	ERICK ESTEBAN	GALLARDO ORTIZ	1998-11-26	QUITO
85a851e09b9069d083d8ef07794a80d7d25520701effcae706710553f369efb6	CRISTHIAN FERNANDO	TOHASA PASPUEZAN	1997-11-30	QUITO
93a628b27d65c3ecce0083a9b43594fed7157e380e02f8ea88ecca6b751882e6	DANNY ESTEBAN	VENEGAS VILLAVICENCIO	1995-03-29	QUITO
3d88e135092259c85b42215b7061498cd669c5ff4424f525150cba9d8006ad0e	FABRICIO	FLORES	2023-02-15	QUITO

Figura 24: Registro para formulario de consulta externa.

[Autor: William Suntaxi]

En la Figura 25, se muestra la información obtenida desde la base de datos.

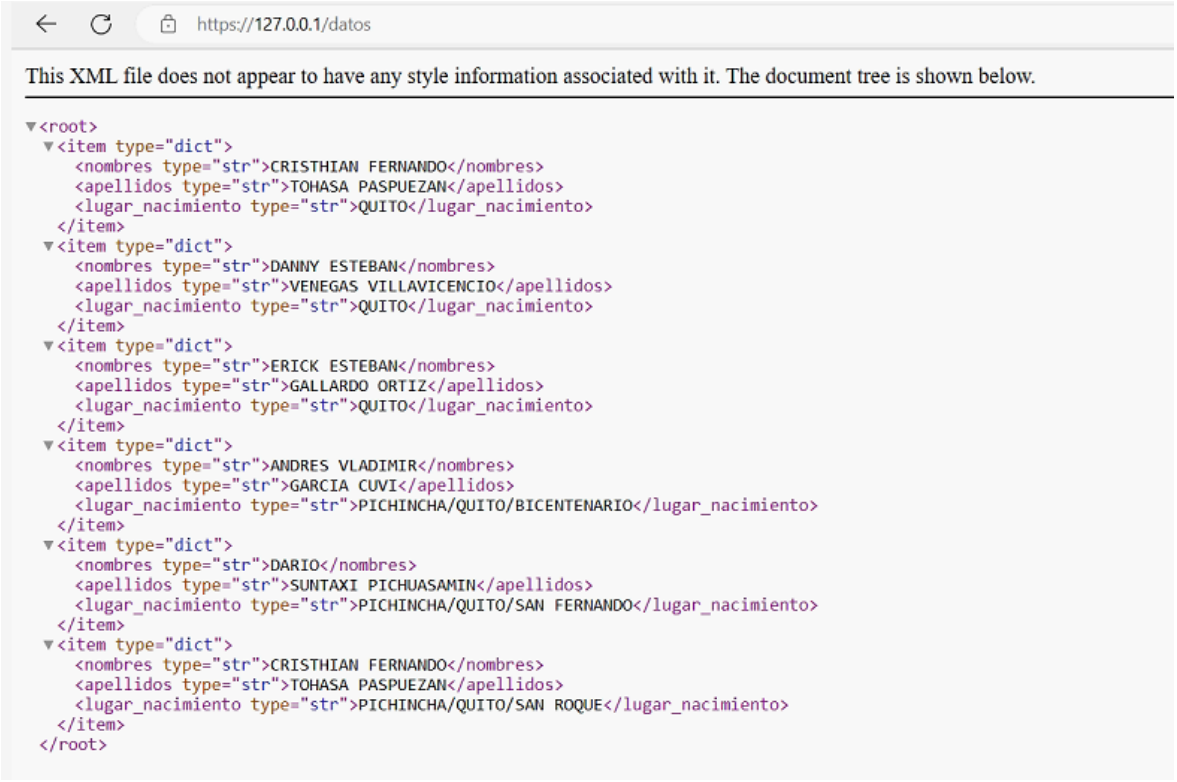

Figura 25: Consulta obtenida de la base de datos.

[Autor: William Suntaxi]

3.1.2 Pruebas de Rendimiento

Las pruebas de rendimiento tienen la finalidad de verificar e identificar si los componentes funcionan de manera eficiente o estas puedan provocar ralentización a la aplicación. Implementar mayor seguridad en una aplicación web puede afectar el rendimiento de un aplicativo. El cifrado de datos puede requerir recursos adicionales del servidor para descifrar los datos, lo que puede aumentar el tiempo de procesamiento y reducir el rendimiento. Por otro lado, la autorización puede requerir mayor tiempo de procesamiento para verificar las credenciales de los usuarios y así permitir el acceso a ciertas áreas de la aplicación.

Asimismo, el aumento de la seguridad puede producir mayor consumo de disco ya que el cifrado de datos ocupa más espacio que los datos sin cifrar, en su ejecución el consumo de disco puede aumentar significativamente. La autenticación mediante tokens genera una gran cantidad de información que debe ser almacenada, de igual forma el almacenamiento en cookies en un aplicativo. Se debe tomar en cuenta que este aplicativo no está almacenando logs del aplicativo que en un ambiente empresarial es muy importante para identificar posibles fallos de seguridad.

Al realizar las validaciones de entradas de usuario, así como también la autenticación y autorización puede requerir un gran consumo de memoria RAM cuando el aplicativo se encuentre en un ambiente de producción con millones de usuarios. De igual forma el cifrado de datos requiere mayor consumo de memoria RAM ya que cada dato aumentará su tamaño y la transferencia de esta información en todo el aplicativo.

Aumentar la seguridad en un aplicativo puede afectar a: el consumo de recursos, a la disponibilidad de la información. Si el aplicativo es demasiado lento, puede suceder que los usuarios no utilicen este aplicativo ya que a nivel corporativo preferirían otras alternativas.

En la Figura 25, se muestra el porcentaje de consumo en disco. Se debe tener en cuenta que cada iteración es cada 10 segundos, en donde el consumo más alto se da en la interacción 43, la cual muestra el mayor consumo en ciclos de disco duro.

Figura 26: Disco físico.

[Autor: William Suntaxi]

En la Figura 27 se tiene la comparación de la escritura de disco en el cual se observa que en el ciclo 45 se da el mayor consumo, esto se debe a que se está realizando la escritura de archivos y la consulta a la base de datos.

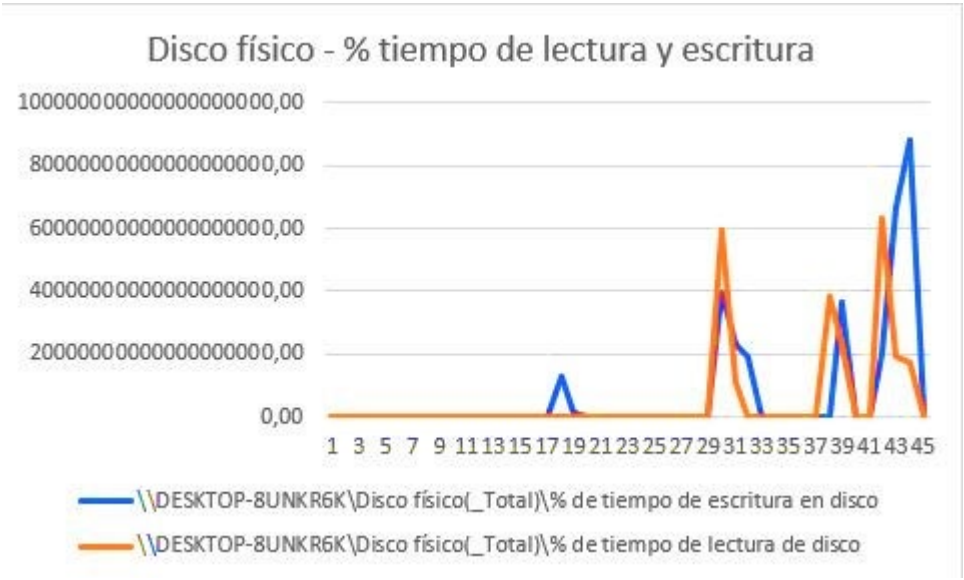

Figura 27: Tiempo de lectura y escritura.

[Autor: William Suntaxi]

En la Figura 28, se muestra un punto bajo de inflexión en el ciclo 3 y en el ciclo 27 se observa una reducción en el consumo de memoria. Esto se debe a que en ese momento la memoria RAM para realizar la escritura en la base de datos y en los archivos.

Figura 28: Memoria en porcentaje de bytes confirmados en uso.

[Autor: William Suntaxi]

En la Figura 29 se observa el consumo de memoria cache, la misma que es irregular. Esta irregularidad se da debido al procesamiento de información y eventos que se están ejecutando de manera paralela en el aplicativo.

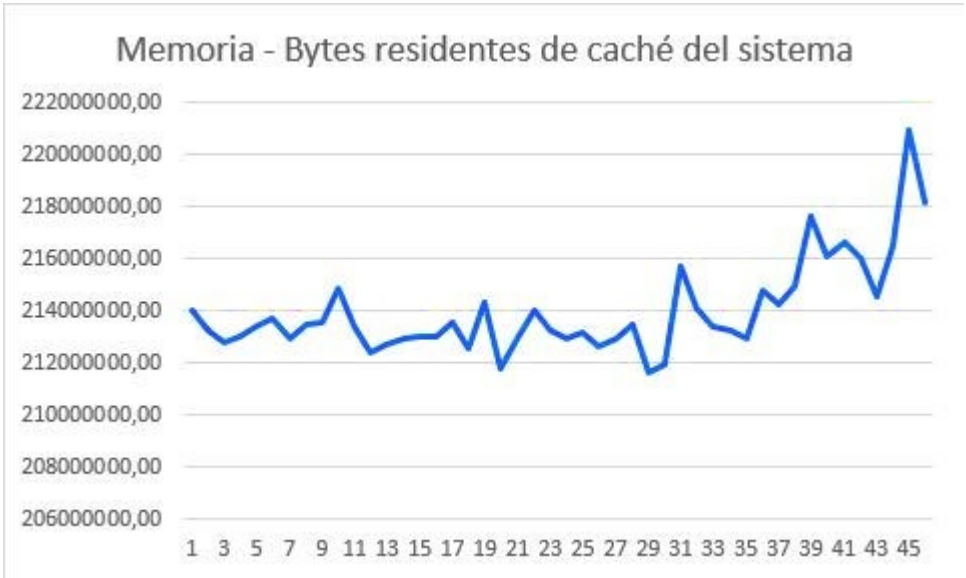

Figura 29: Consumo de memoria RAM.

[Autor: William Suntaxi]

3.1.3 Pruebas de Resiliencia

Usando la herramienta Microsoft Threat Modeling Tool, se ingresó la arquitectura vista en la figura 15 la cual permitió obtener un reporte (Véase Anexo III) detallado de las amenazas que nuestra arquitectura pudiera recibir.

3.2 Discusión de Resultados

Se realizó unas pruebas del flujo del sistema recorriendo cada módulo y registrando el rendimiento del computador tomando en cuenta el procesador, la memoria y el disco. Esto se realizó con iteraciones de 20 segundos. Es decir, cada 20 segundos se toma una muestra del rendimiento y se lo guarda en un archivo con extensión *.csv.

Como se puede observar en las Figuras 26, 27, 28 y 29, los primeros 5 minutos del flujo no generó mucho movimiento. Esto quiere decir que, al momento de llenar el formulario, validar los datos, aceptar los términos y condiciones no hubo una carga alta en el rendimiento, pero al momento de terminar la transferencia y realizar la inserción de los datos en la base de datos el rendimiento aumentó, especialmente en la escritura en el disco alrededor de la iteración 15, como se muestra en la Figura 27.

Para finalizar al final del flujo del sistema se realiza la consulta a la base y se retorna un XML para ser mostrado al usuario que fue autorizado anteriormente.

3.3 Conclusiones

Al finalizar el proyecto se pudo obtener las siguientes conclusiones:

- Los requerimientos funcionales identificados e implementados en esta investigación durante la recolección y procesamiento fueron los siguientes: *Autenticación*, se realizó la implementación de un Login usando certificados digitales, de igual forma por la parte de enmascaramiento de datos se implementó una función para utilizar una técnica de Hashing más salting en la información crítica de los clientes para evitar las colisiones en los datos. Finalmente se implementó una base de datos temporal para el acceso a la información para entidades externas.
- Los requerimientos no funcionales identificados e implementados fueron las técnicas de Hashing para el ocultamiento de la información, además se implementó un sistema web en el cual se garantizó que la información esté disponible para los usuarios. Finalmente se realizó un análisis de rendimiento para comprobar el consumo de memoria RAM, DISCO, RED y PROCESADOR.
- Se realizó de manera exitosa la implementación de un aplicativo web considerando como referencia el GDPR europeo, la LOPDP ecuatoriana y el esquema Gubernamental de seguridad de información EGSI. Para lo cual se tomaron las siguientes consideraciones en el desarrollo: identificación y validación de los datos

personales, ocultamiento de la información personal por medio de las técnicas de minimización y pseudonimización de datos. Así como la realización de la evaluación de riesgos asociados a la recolección de datos personales.

- Al realizar las pruebas de flujo en un entorno controlado, para el análisis de consumo de recursos del procesador, memoria RAM, caché y disco, no se encontraron comportamientos anormales en el aplicativo implementado. Se puede observar que, al realizar la persistencia de datos, estos tienen un consumo mayor de recursos. Puesto que, se están utilizando recursos adicionales para el ocultamiento de información. Así como también para la transferencia segura de la información por los diferentes módulos hasta llegar al módulo de almacenamiento.
- Uno de los principales inconvenientes de tener una elevada seguridad es la afectación a los recursos computacionales, esto provoca que no funcionen adecuadamente, y se ralentice la capacidad de respuesta. Además, la experiencia de usuario se ve afectada, ya que, se debe realizar demasiados pasos de autenticación. Lo que puede afectar a la accesibilidad. Por otro lado, tener demasiada seguridad también puede llevar a fallos de seguridad, esto se debe a que, durante la implementación, la complejidad aumenta y un fallo de seguridad puede ser difícil de detectar por lo que se pueden dejar posibles vulnerabilidades. Por lo tanto, es importante realizar un correcto análisis de riesgo y amenazas, así como también un diseño óptimo en el desarrollo.

3.4 Recomendaciones

- Para la correcta realización de un modelo de amenazas es importante identificar los activos críticos del sistema, esto incluye la información de identificación personal. Además, se debe identificar las amenazas potenciales para luego evaluar el impacto de estas. Seguido de este proceso se debe identificar las vulnerabilidades
- Identifica los activos críticos del sistema PII: Comienza identificando los datos y activos críticos del sistema PII. Esto incluye información de identificación personal, como nombres, números de identificación, direcciones, información financiera, etc.

  Identifica las amenazas potenciales: Identifica las amenazas potenciales que pueden afectar los activos críticos. Algunas posibles amenazas incluyen ataques de hackers, amenazas internas, errores humanos, desastres naturales, etc.

Evalúa el impacto de las amenazas: Evalúa el impacto que cada amenaza podría tener en los activos críticos del sistema PII. Esto te ayudará a priorizar y enfocarte en las amenazas más críticas.

Identifica las vulnerabilidades: Identifica las vulnerabilidades que existen en el sistema PII que podrían ser explotadas por las amenazas. Algunas vulnerabilidades comunes incluyen contraseñas débiles, falta de parches de seguridad, configuraciones incorrectas, etc.

Prioriza las amenazas y vulnerabilidades: Prioriza las amenazas y vulnerabilidades según su impacto y probabilidad de ocurrencia.

Desarrolla contramedidas: Desarrolla contramedidas para cada amenaza y vulnerabilidad identificada. Esto puede incluir implementar medidas de seguridad técnicas, políticas y procedimientos de seguridad, capacitación de empleados, entre otras.

- Realiza pruebas y evaluaciones periódicas: Realiza pruebas y evaluaciones periódicas para garantizar que el modelo de amenazas siga siendo relevante y efectivo a medida que cambien las amenazas y vulnerabilidades.
- Es importante evaluar la probabilidad de que ocurra una amenaza y el impacto que tendría en los activos críticos, esto permitirá priorizar la mitigación de riesgos.
- La seguridad se debe tomar en cuenta en todo el proceso de desarrollo, desde el diseño, hasta la implementación del software, en este caso el aplicativo web. A su vez, se debe tener en cuenta que el rendimiento de un aplicativo es inversamente proporcional a la seguridad que se le aplica.
- Es importante saber que la recolección de datos en un aplicativo para el posterior análisis es crítica en la investigación. Puesto que, se debe definir un objetivo de estudio, esto ayudará a entender que información se debe recolectar y que información es crítica o confidencial para una organización.

4. REFERENCIAS BIBLIOGRÁFICAS

[1] A. N. R. DEL ECUADOR, “LEY ORGÁNICA DE PROTECCIÓN DE DATOS PERSONALES,” 2021. https://www.registrospublicos.gob.ec/programas-servicios/servicios/proyecto-de-ley-de-proteccion-de-datos/

[2] Naranjo, Martìnez, and Subìa, “Entra en vigencia la Ley Orgánica de Protección de Datos Personales,” 2021. https://nmslaw.com.ec/ley-organica-proteccion-datos-

personales/ (accessed Aug. 18, 2022).

[3] U. Europea, “General Data Protection Regulation,” 2018. https://gdpr-info.eu/

[4] Y. Europe, “Reglamento general de protección de datos,” 2018. https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_es.htm

[5] M. B. Samper, “Reglamento General De Protección De Datos (UE),” *Protección de datos personales*, 2020. https://www.boe.es/doue/2016/119/L00001-00088.pdf

[6] MINISTERIO DE TELECOMUNICACIONES Y DE LA SOCIEDAD DE LA INFORMACIÓN, “ESQUEMA GUBERNAMENTAL DE SEGURIDAD DE LA INFORMACIÓN (EGSI),” 2020. https://www.gobiernoelectronico.gob.ec/egsi/

[7] Suments, “Auge de leyes nacionales sobre protección de datos.” https://suments.com/es/tendencias-proteccion-de-datos-2022/

[8] NCSL, “2022 Consumer Privacy Legislation,” 2022. https://www.ncsl.org/research/telecommunications-and-information-technology/2022-consumer-privacy-legislation.aspx

[9] ATICO34, “Nueva Ley de protección de datos en China o PIPL,” 2021. https://protecciondatos-lopd.com/empresas/ley-china-pipl/

[10] N. Autoridad and P. de D. (ANPD, “LGPD – Ley General de Protección de Datos de Brasil,” 2020. https://www.cookiebot.com/es/lgpd/

[11] Agencia Española de Protección de Datos, *Guía de Privacidad desde el Diseño*. 2019. [Online]. Available: https://www.aepd.es/sites/default/files/2019-11/guia-privacidad-desde-diseno.pdf

[12] Ayuda ley proteccion datos, “Ejemplos de seudonimización de datos,” 2023. https://ayudaleyprotecciondatos.es/2020/11/18/seudonimizacion-de-datos/ (accessed Mar. 01, 2023).

[13] G. Atico, “Tipos de seudonimización,” 2022. https://protecciondatos-lopd.com/empresas/seudonimizacion-anonimizacion/ (accessed Mar. 01, 2023).

[14] Ecityclic, “Lo que debes saber sobre la Seudonimización de los datos,” 2020. https://opensistemas.com/seudonimizacion-y-anonimizacion-de-datos/

[15] S. B. SANTOS DIVINO, “Reflexiones escépticas, principiológicas y económicas sobre el consentimiento necesario para la recolección y tratamiento de datos.,” pp.

179–206, 2019, doi: http://dx.doi.org/10.18800/derechopucp.201902.006.

[16] F. N. Solarte-Solarte, Edgar Rodrigo Enriquez Rosero, and M. del C. Benavides, “Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma ISO/IEC 27001,” *Rev. Tecnológica - ESPOL*, vol. 28, no. 5, 2015, [Online]. Available: http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/456

[17] F. N. Roldán Carrillo, “Los ejes centrales de la protección de datos: consentimiento y finalidad. Críticas y propuestas hacia una regulación de la protección de datos personales en Ecuador,” *USFQ Law Rev.*, vol. 8, no. 1 SE-Artículos, pp. 175–202, May 2021, doi: 10.18272/ulr.v8i1.2184.

[18] A. Hevner *et al.*, “Design Science in Information Systems Research,” *Manag. Inf. Syst. Q.*, vol. 28, pp. 75-, 2004.

[19] S. B. Santos Divino, “Reflexiones escépticas, principiológicas y económicas sobre el consentimiento necesario para la recolección y tratamiento de datos,” *Derecho PUCP*, no. 83, pp. 179–206, 2019, doi: 10.18800/derechopucp.201902.006.

[20] E. P. Nacional, “POLÍTICA DE USO DE LA INFORMACIÓN, ACTIVOS DE INFORMACIÓN INSTITUCIONAL Y SEGURIDAD INFORMÁTICA,” Quito, 2018. [Online]. Available: https://www.epn.edu.ec/wp-content/uploads/2021/05/Proyecto-Política-de-la-Información.pdf

[21] PAOLA UTRERAS, “GESTIÓN DE IDENTIDAD DIGITAL DE USUARIOS EN SERVICIOS WEB PARA LA PROTECCIÓN DE LA PRIVACIDAD DE LA INFORMACIÓN,” PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR ESMERALDAS, 2021. [Online]. Available: https://repositorio.pucese.edu.ec/bitstream/123456789/2419/1/UTRERAS LOGACHO PAOLA LISSETTE.pdf

[22] F. N. J. Solarte-Solarte, E. R. Enriquez-Rosero, and M. del C. Benavides-Ruano, “Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma ISO/IEC 27001.,” *Rev. Tecnológica - ESPOL*, vol. 28, no. 5, pp. 497–498, 2015, [Online]. Available: http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/456/321

[23] V. N. E. Sebastián María Alejandra, “MODELO DE PREVENCIÓN PARA EL TRATAMIENTO DE DATOS PERSONALES.,” La Plata, 2021. [Online]. Available: http://sedici.unlp.edu.ar/bitstream/handle/10915/128821/Documento_completo.Seb

astian-N.E.Vazquez.pdf-PDFA.pdf?sequence=1&isAllowed=y

[24] S. V. Echeverría Rodríguez, “Evaluación del sistema de gestión de la seguridad de la información de la Empresa Eléctrica Quito utilizando la norma ISO 27002,” 2015, [Online]. Available: http://bibdigital.epn.edu.ec/handle/15000/12000

[25] C. E. Cáceres Tarco and C. E. Mena González, “Elaboración de la guía de implantación de las normas prioritarias del esquema gubernamental de seguridad de la información EGSI en las entidades de la administración pública central,” 2015, [Online]. Available: http://bibdigital.epn.edu.ec/handle/15000/11234

[26] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” *J. Manag. Inf. Syst.*, vol. 24, pp. 45–77, 2007.

[27] I. Williams and X. Yuan, “Evaluating the Effectiveness of Microsoft Threat Modeling Tool,” 2015. doi: 10.1145/2885990.2885999.

5. ANEXOS

En esta sección se encuentran los documentos pertinentes que apoyaron la realización del proyecto y estos documentos son:

ANEXO I. Estado del Arte

ANEXO II. Documentos Scrum

ANEXO III. Reporte de Amenazas

ANEXO IV. Evaluación de Riesgos

ANEXO V. Video demostrativo

ANEXO VI. Código fuente de la aplicación

ANEXO I

En el siguiente enlace se puede encontrar el documento inicial. Un estado de arte realizado por el autor del proyecto que tiene como objetivos ayudar con las bases teóricas y prácticas para la realización de la tesis.

Enlace:

https://epnecuador-my.sharepoint.com/:f:/g/personal/willian_suntaxi_epn_edu_ec/EoUo6qbvpNxKoaGxN6WSe_UB1lRGk4FrbsxbzF6DUSuCqg?e=kf3SlR

ANEXO II

En el siguiente enlace se puede encontrar los documentos obtenidos al realizar Scrum. Entre ellos se encuentran las historias de usuario, los Sprints y la calendarización del proyecto.

Enlace:

https://epnecuador-my.sharepoint.com/:f:/g/personal/willian_suntaxi_epn_edu_ec/EsSP_Nr3M2hHisFEMd3gVdIB682yIVGEmbyQ5oCmukJAag?e=1h8qdM

ANEXO III

En el siguiente enlace se puede encontrar el documento que dio de resultado la herramienta de modelado de amenazas. Un documento que contiene las 76 amenazas que fueron visualizadas por la herramienta

Enlace:

https://epnecuador-my.sharepoint.com/:f:/g/personal/willian_suntaxi_epn_edu_ec/Eolxy2swOHRAv0i7cca9jIIB8cXgSqUpZ01BTfylgdMDcw?e=Ocpddo

ANEXO IV

En el siguiente enlace se puede encontrar el documento con el resultado del análisis de riegos. Un documento que contiene las amenazas y las contingencias del componente de Recolección de datos.

Enlace:

https://epnecuador-my.sharepoint.com/:f:/g/personal/willian_suntaxi_epn_edu_ec/ElqrHSgyoDpDoIx8RMhbD7QBpgh4pdm_4WjTEqzSXHU2Vw?e=vXK2WW

ANEXO V

En el siguiente enlace se puede encontrar el video que muestra el funcionamiento del prototipo. Aquel que fue obtenido al realizar el proyecto de integración curricular (TIC).

Enlace:

https://epnecuador-my.sharepoint.com/:f:/g/personal/willian_suntaxi_epn_edu_ec/EpweDqaa9uVCqh0-oVBgBc4B0YTndKz4uitesXNRB7dFgw?e=gHQR5u

ANEXO VI

En el siguiente enlace se puede encontrar el video que muestra el funcionamiento del prototipo. Aquel que fue obtenido al realizar el proyecto de integración curricular (TIC).

Enlace:

En el siguiente enlace se puede encontrar el código fuente de la aplicación integrada de los tres componentes. Aquel que fue obtenido al realizar el proyecto de integración curricular (TIC).

Enlace: https://github.com/WILLIAMSUNTAXI/Tesis_2023