

Universidad Siglo 21



Trabajo Final de Grado. Trabajo de Prototipado Tecnológico.

Carrera: Licenciatura en Informática

Sistema de Gestión de Documentos Digitales con Almacenamiento

Descentralizado

Autor: Mario Andrés Bedin

Legajo: VINF-06998

Rosario, Julio 2021

Índice

Resumen	4
Abstract.....	5
Título	6
Introducción.....	6
Antecedentes.....	6
Descripción del área problemática	7
Justificación	7
Objetivo General del Proyecto	8
Objetivo Especifico del Proyecto	9
Marco Teórico Referencial.....	9
Dominio del Problema.....	9
TICs	10
Competencia	11
Diseño Metodológico	13
Metodología.....	13
Herramientas.....	14
Relevamiento	16
Relevamiento Estructural	17
Relevamiento Funcional.....	18
Procesos de Negocio.....	21
Diagnóstico y Propuesta	23
Diagnóstico.....	23
Propuesta	25
Objetivo, Límite y Alcances del Prototipo	26
Objetivo del Prototipo	26
Límites	26
Alcances.....	26
Descripción del sistema	27
Product Backlog	27
Historias de Usuario	27
Estructuras de Datos	32
Prototipos de Interfaces de Pantallas	33

Diagrama de Arquitectura	36
Seguridad	36
Acceso a la Aplicación	36
Políticas de respaldo	38
Análisis de Costos	38
Hardware	39
Software.....	40
Recursos Humanos	41
Análisis de Riesgos.....	42
Tabla de riesgos	42
Clasificación de riesgos	44
Conclusiones.....	47
Demo	48
Referencias	49
Anexos	51
Anexo I - Leyes	51
Anexo II - Significado de acrónimos y códigos utilizados.....	51
Anexo III - Comparación de costos de servicios de almacenamiento.....	52

Índice de Ilustraciones

Ilustración 1- Tabla Gantt – Análisis y Documentación	15
Ilustración 2 - Tabla Gantt – Desarrollo y Pruebas	15
Ilustración 3 - Diagrama de Gantt – Análisis y Documentación.....	16
Ilustración 4 - Diagrama de Gantt – Desarrollo y Pruebas.....	16
Ilustración 5 - Organigrama Empresa Modelo	18
Ilustración 6 - Diagrama de flujo - Almacenar y Recuperar Documento.....	22
Ilustración 7 - Product Backlog.....	27
Ilustración 8 - Sprint Backlog - parte 1	30
Ilustración 9 - Sprint Backlog - parte 2	31
Ilustración 10 - Sprint Backlog - parte 3	31
Ilustración 11 - Sprint Backlog - parte 4	31
Ilustración 12 - Sprint Backlog - parte 5	32

Ilustración 13 - Estructura de Datos	32
Ilustración 14 - Pantalla Login	33
Ilustración 15 - Pantalla Registro de Usuario	33
Ilustración 16 - Pantalla Upload Documento	34
Ilustración 17 - Pantalla Listar Documentos	34
Ilustración 18 - Lista de archivos en Nodo IPFS (Filecoin Hot Layer).....	34
Ilustración 19 - Reporte generado mediante analítica de datos.	35
Ilustración 20 - Diagrama de Arquitectura	36
Ilustración 21 - Comparativa de Costos Filecoin vs S3.....	52

Índice de Tablas

Tabla 1- Proceso Almacenar constancia de cobertura.....	23
Tabla 2- Proceso recuperar constancia de cobertura	24
Tabla 3 - Proceso visualizar reporte de documentos almacenados	24
Tabla 4 - Registro de Usuario.....	27
Tabla 5 - Login de Usuario.....	28
Tabla 6 - Upload de Documento.....	28
Tabla 7 - Listar Documentos	29
Tabla 8 - Descarga de Documento	29
Tabla 9 - Generar Reporte	30
Tabla 10 - Costos de Hardware	39
Tabla 11 - Costos de Software.....	40
Tabla 12 - Costos de Recursos Humanos	41
Tabla 13 - Costos Totales	41
Tabla 14 - Listado de Riesgos.	42
Tabla 15 - Clasificación de Riesgos	44

Resumen

Actualmente la gestión de documentos digitales es una tarea clave para muchas organizaciones, tanto del ámbito público como privado. Especialmente en los casos que dicha documentación representa información sensible que debe ser preservada de manera segura, fiable y en donde se garantice en todo momento la integridad de la misma.

Se pudo observar, que cuando los volúmenes de documentación aumentan considerablemente, preservar la fiabilidad e integridad de la misma se convierte en una tarea sumamente compleja y costosa que requiere una infraestructura segura y escalable. Esta observación permitió pensar en la tecnología de cadena de bloques como principal herramienta para brindar soporte a estos requisitos, por la naturaleza de la misma y por la posibilidad de crear contratos inteligentes para gestionar el acceso a la misma.

Esta necesidad fue detectada principalmente mediante observación directa trabajando por más de cuatro años en múltiples implementaciones de empresas del rubro asegurador argentino y desempeñándome como consultor de la plataforma Guidewire tanto en Argentina como Reino Unido durante los últimos tres años.

Otro de los aspectos importantes fue el de poder generar información de valor mediante procesos de reportes, utilizando analítica de datos, lo cual brinda una herramienta importante para la toma de decisiones relacionadas a la gestión de los documentos administrados.

Palabras claves: Gestión de documentos, Cadena de Bloques, Analítica de datos, Seguridad

Abstract

Nowadays the document management system is a key process for every public or private organizations. Especially in cases where such documents represent sensitive information that must be preserved in a safe and reliable manner. Also, where its integrity needs to be guaranteed for a long time.

It has been noted that when the amount of documentation increases considerably, preserving its reliability and integrity becomes an extremely complex and expensive task that requires a secure and scalable infrastructure. This observation let us think in blockchain technology as a main tool to support these requirements, due its nature and the possibility of implement smart contracts to manage the resources' access.

This need was detected mainly by direct observation, working for more than four years in multiple Argentinian's insurer core implementations and serving as a Guidewire platform consultant in both Argentina and the United Kingdom during the last three years.

Another important goal achieved was the ability to generate valuable information through reporting processes, using Data Analytics, which provides an important tool to support business managers' decisions related document management process.

Keywords: Document management, Blockchain, Data analytics, Security

Título

Sistema de gestión de documentos digitales con almacenamiento descentralizado.

Introducción

El almacenamiento y administración de documentación sensible es una de las tareas más importantes que se llevan a cabo dentro de las organizaciones, sean estas públicas o privadas, en muchos casos es de vital importancia preservar dichos documentos de manera segura, garantizando que solo las personas con el correcto nivel de permisos tengan accesos a los mismos.

El presente proyecto tiene por finalidad mejorar el proceso de gestión de documentación digital dentro de las organizaciones, mediante el uso de almacenamiento descentralizado soportado por la tecnología blockchain. Asegurando la integridad de la información durante todo el ciclo de vida del documento.

Antecedentes

La problemática a desarrollar no es nueva, ya en 1880 Edwin Seibels para dar una primera solución inventó lo que todos conocemos como gabinete archivador o por su nombre en inglés 'file cabinet', con esto se intentó poner un poco de orden en un mundo con un serio problema de gestión de documentos. Luego con los avances tecnológicos e informáticos y la aparición de los primeros servidores y PC's que conectadas a una red LAN lograban estar distribuidas, las empresas pudieron crear y administrar documentos de manera digital, lo cual trajo muchos beneficios como la administración electrónica, reducción del uso de papel. Pero también comenzaron a aparecer nuevas problemáticas como la dificultad para rastrear, la seguridad, y dificultad para administrar los documentos.

Con el correr de los años, entre la década del 80 y 90, comenzaron a aparecer los primeros EDMS (Electronic Document Management System) que introdujeron importantes mejoras como: colaboración entre clientes, manejo de accesos por roles, captura de datos en múltiples formatos, destrucción automatizada de archivos (archiving). De igual manera con este nuevo cambio, aparecieron nuevas problemáticas relacionadas

a la alineación de los EDMS con los procesos comerciales, problemas en la búsqueda de documentos (Document Logistix, 2021).

En la actualidad con el avance de internet, y el gran incremento de servicios en la nube, las empresas no dependen de sus limitaciones de infraestructura local para atacar la problemática, sino que pueden elegir en base a sus posibilidades y necesidades servicios y productos en la nube, pueden contratar soluciones del tipo SAAS (Software as a Service) con lo cual pueden tener el DMS pre construido en la nube, y el software bajo demanda desde un navegador. El principal beneficio con esto es que hay muchas soluciones ajustadas a las necesidades y la capacidad que tiene la empresa, estas soluciones permiten múltiples formas de colaboración remota.

Descripción del área problemática

La problemática radica en que con el correr de los años la tarea de decidir cual es la opción adecuada de DMS para las empresas u organizaciones es cada vez es mas compleja, que a pesar de contar con muchas opciones los precios han ido incrementándose con el correr de los años, y muchas empresas no cuentan con los recursos para contratar una solución adecuada para administrar su documentación.

Esto se intensifica en el sector público, donde claramente los recursos económicos en muchas dependencias son escasos, y deben estar abocados a solucionar otro tipo de problemáticas mas urgentes. A su vez las soluciones actuales cuentan con almacenamiento centralizado de los datos y esto mismo hace que puedan aparecer fallas de seguridad en cuanto a su acceso y/o almacenamiento.

Justificación

Para mitigar la problemática descrita en el párrafo anterior se desarrolló un sistema que mejora notablemente los costos de servicio de almacenamiento y administración de la documentación, reduce las fallas de seguridad que tienen los sistemas actuales incorporando blockchain como componente central de la solución. Esto va a permitir que muchas organizaciones que no tenían acceso a este tipo de soluciones, puedan mejorar los procesos de negocio impactados por la gestión de documentos.

Cabe mencionar que en la actualidad existen muchas soluciones similares que atacan múltiples problemáticas, pero solo unas pocas proveen almacenamiento descentralizado mediante el uso de redes blockchain, para afirmar esto tomo como fuente de información los estudios realizados por TI Gartner, que son utilizados mundialmente como referencia en el sector, se puede observar en el cuadrante mágico relacionado a las aplicaciones de tipo DMS (Gartner Magic Quadrant, 2021) no se ha observado siquiera una solución que utilice blockchain para mitigar la problemática mencionada.

Dentro de los beneficios que se han conseguido con la implementación de la solución se enumeran:

- Reducción notable del costo de almacenamiento de información. En base a precios de Julio de 2021, tomando como parámetro la blockchain Filecoin, el costo por GB por mes, de una solución es solo el 0.11% del costo de Amazon S3. (Filecoin Miner Analytics, 2021). Se pueden observar los cálculos pertinentes la sección de anexos.
- Mejorar en la seguridad, accesibilidad de la información almacenada. Por medio de un Smart contract, se va a garantizar que solo las personas con los roles adecuados puedan acceder a la información solicitada.
- Mejorar la integridad de la información: Este beneficio se logra por la naturaleza propia de almacenar información de manera descentralizada en una red blockchain.
- Ofrecer información estadística importante relacionada al uso que se le da a la documentación en la aplicación por medio de analítica de datos para dar soporte a la toma de decisiones de la organización.

Objetivo General del Proyecto

Diseñar y desarrollar un sistema web que permita administrar documentación sensible de empresas privadas u organizaciones públicas brindando almacenamiento descentralizado vía blockchain a muy bajo costo y generar información relevante que ayude a mejorar procesos de negocio garantizando en todo momento la integridad de la información.

Objetivo Especifico del Proyecto

- Analizar los componentes principales de un sistema de gestión de documentos.
- Analizar y comparar blockchains orientadas a almacenamiento distribuido de archivos para integrar al sistema.
- Diseñar y Desarrollar aplicación web que permita almacenar y recuperar documentación electrónica desde una blockchain.
- Diseñar y Desarrollar Smart contract que valide la autorización de accesos a documentos almacenados de manera distribuida para resguardar la integridad de los mismos.
- Identificar indicadores importantes de un sistema de gestión de documentos que puedan ayudar en la toma de decisiones para mejorar procesos de negocios.

Marco Teórico Referencial

Dominio del Problema

Un DMS (Document Management System) es una de las herramientas claves en toda organización o empresa. Dentro de las actividades principales que abarca se encuentran almacenar, administrar, y controlar el flujo de documentos dentro de la organización asegurando la integridad de la información y el correcto acceso a esta. Adicionalmente también debe proveer mecanismos de archivado o autodestrucción, configuración de flujos de trabajo (comúnmente llamados Workflows) que a alto nivel es el movimiento automatizado de documentos en respuesta a acciones que suceden en los distintos procesos de negocio.

En la actualidad estas soluciones están disponibles para todos los tipos de organizaciones y/o empresas, pero con el correr del tiempo y los avances tecnológicos a veces es complejo seleccionar la herramienta adecuada. Las herramientas más populares se pueden contratar como servicios de almacenamiento en la nube y en muchos casos cuando la cantidad de documentos, y las interacciones usuario-sistema comienza a escalar el costo de mantenimiento se vuelve demasiado costoso para muchas organizaciones. Desde el marco normativo se han producido grandes avances que han ayudado a impulsar el almacenamiento digital de la documentación y potencian continuamente este tipo de

sistemas, se pueden mencionar para el caso particular de Argentina la promulgación de la ley 25506 de 2001 por la cual se da validez a la firma digital y se presenta un marco normativo para regular los documentos generados en formato digital (Poder legislativo Nacional, 2001). En el ámbito público, también tuvo mucho impacto la ley 27446 que tiene como principal objetivo la simplificación y desburocratización de la administración pública (Poder Legislativo Nacional, 2018).

En cuanto a la protección de datos, todas las herramientas actuales brindan cierto nivel de seguridad, pero esto puede ser superado notablemente con la irrupción de la tecnología blockchain aplicada a almacenamiento de documentos. Hay estudios que fundamentan esta última afirmación, como el de la consultora Capterra (Capterra, 2018) en el que se destaca que con dicha aplicación se lograría ahorrar mucho en costos de mantenimiento, el sistema sería muy difícil de hackear y se ganaría flexibilidad por medio de la posibilidad de implementar smart contracts para automatizar ciertas transacciones e interacciones.

TICs

A continuación, se describen las tecnologías utilizadas en el desarrollo del sistema.

- **Javascript:** lenguaje de programación que funciona en los navegadores de forma nativa (lenguaje interpretado sin necesidad de compilación). Se utiliza como complemento de HTML y CSS para crear páginas webs. Permite realizar llamadas JSON RPC para la interacción con servicios web. (Web oficial javascript, 2021)
- **NodeJS:** Es un entorno JavaScript que nos permite ejecutar en el servidor, de manera asíncrona, con una arquitectura orientada a eventos y basado en el motor V8 de Google, dicho compila Javascript en código máquina nativo en vez de interpretarlo en el navegador, consiguiendo así una velocidad mucho más alta. Node es de código abierto y puede ejecutarse en Mac OS X, Windows y Linux. (Web oficial Node.js, 2021)
- **Bootstrap:** Framework de diseño web que permite desarrollar interfaces web con CSS y Javascript con la particularidad que permite realizar soluciones del tipo responsive, lo que hace que la interfaz web se adapte al dispositivo desde donde se esta accediendo al sitio. (Sitio oficial bootstrap, 2021)

- **Git:** Un sistema de control de versiones nos va a servir para trabajar en equipo de una manera mucho más simple y óptima cuando estamos desarrollando software. Con la principal ventaja que no necesitamos estar conectados todo el tiempo al repositorio remoto para poder generar nuevas versiones de nuestro código. (Sitio oficial Git, 2021)
- **Gitflow:** Es un flujo de trabajo pre establecido que garantiza el correcto despliegue de versiones de nuestro sistema, facilitando la integridad al momento de incorporar nuevas funcionalidades y/o corrección de errores (Sitio oficial Atlassian, 2021).
- **MongoDB:** Es un sistema de base de datos NoSQL, orientado a documentos y de código abierto. En lugar de guardar los datos en tablas, tal y como se hace en las bases de datos relacionales, permite guardar estructuras de datos BSON (una especificación similar a JSON) con un esquema dinámico, haciendo que la integración de los datos en ciertas aplicaciones sea más fácil y rápida. (Sitio oficial MongoDB, 2021)
- **Filecoin:** Es una blockchain que funciona bajo una prueba novedosa, llamada Prueba de Espacio Tiempo (Proof-of-Spacetime o PoST), donde los bloques son creados por mineros que están almacenando datos. El protocolo de Filecoin proporciona un servicio de almacenamiento y recuperación de datos a través de una red de proveedores de almacenamiento independientes que no depende de un solo coordinador, donde: Los clientes pagan para almacenar y recuperar datos. Los mineros ganan tokens ofreciendo almacenamiento, es decir, los mineros ganan tokens mediante la entrega de datos. (Sitio oficial Filecoin, 2021).

Competencia

En la actualidad existen múltiples soluciones que pueden ser competencia del sistema desarrollado, la mayoría de ellas son de las llamadas tradicionales y solo unas pocas utilizan blockchain como agregado de valor e innovación. A continuación, se detalla alguna de las principales:

- **OpenText:** Las aplicaciones de OpenText administran contenido o datos no estructurados para la administración, eficacia y monetización requeridas en compañías grandes, agencias de gobierno y empresas de servicio profesional. Las soluciones de OpenText pretenden cumplir los requisitos de administración de la

información, incluyendo la administración de volúmenes grandes de datos, conformidad con requisitos reguladores, y experiencia móvil y administración en línea. Además del almacenamiento Opentext provee otras funcionalidades como la creación, revisión, aprobación, gestión de documentos mediante propiedades pre establecidas. (Sitio oficial Opentext, 2021).

- **Box:** Es una solución de intercambio de archivos basadas en la nube destinadas a brindar a sus clientes la libertad de ver, editar y colaborar en documentos cuando y donde lo necesiten. Ofrece diferentes niveles de soluciones comerciales, que cubre los requisitos básicos, avanzados y empresariales, que están destinados a satisfacer todas las partes del mercado corporativo con diferentes necesidades y presupuestos. (Sitio oficial Box, 2021).
- **Alfresco:** Esta herramienta ofrece capacidades de gestión de contenido empresarial (ECM) abiertas, flexibles y sumamente escalables. El contenido es accesible desde cualquier lugar y de la forma en la que usted trabaje, independientemente de la ubicación, los repositorios o los dispositivos, y se integra fácilmente con otras aplicaciones. (Sitio oficial Alfresco, 2021).
- **Amazon S3:** Amazon Simple Storage Service (Amazon S3) es un servicio de almacenamiento de objetos que ofrece escalabilidad, disponibilidad de datos, seguridad y rendimiento líderes en el sector. Gracias a Amazon S3, clientes de todos los tipos y sectores pueden almacenar y proteger cualquier volumen de datos para los más variados fines, como usarlos en lagos de datos, sitios web, aplicaciones móviles, procesos de copia de seguridad y restauración, operaciones de archivado, aplicaciones empresariales, dispositivos IoT y análisis de big data. Amazon S3 proporciona características de administración fáciles de utilizar que le permiten organizar los datos y configurar sofisticados controles de acceso con objeto de satisfacer sus requisitos empresariales, organizativos y de conformidad. (Sitio oficial Amazon, 2021).
- **DocFlow:** Es un software de gestión de documentos basado en blockchain que digitaliza todo el ciclo del papeleo y utiliza mecanismos de contratos inteligentes avanzados para garantizar la autenticidad y seguridad de los datos sin concesiones. También provee una interfaz de usuario web para interactuar con la plataforma

como administrador, API Gateway para acceso de datos, almacenamiento descentralizado. (Sitio oficial PixelPlex, 2021).

Diseño Metodológico

Metodología

Para organizar las distintas etapas del desarrollo se utilizarán las siguientes metodologías y herramientas:

- **Scrum:** Marco de trabajo definido dentro de las metodologías ágiles que se caracterizan por su flexibilidad y capacidad de modificar el producto a lo largo del proyecto bajo un desarrollo iterativo e incremental, donde los requisitos y soluciones evolucionan con el tiempo según la necesidad del proyecto. Esta metodología divide el proyecto en fases (sprints) en las cuales se busca agregar funcionalidades definidas al producto mínimo viable (PMV). (Sitio oficial Scrum, 2021)
- **TDD (Test Driven Development):** Es una técnica que cambia el orden establecido en cuanto a primero desarrollar (programar) y luego probar, de manera que primero se define las pruebas (casos de prueba) y a partir de estos se va desarrollando la funcionalidad, repitiendo el ciclo. A medida que se van añadiendo nuevas funcionalidades, se añaden nuevas pruebas, las cuales (naturalmente) fallan al no estar implementadas, luego se desarrolla el software necesario para aprobar el caso de prueba, hasta que todos tengan éxito. Es una de las técnicas de pruebas de software que más está alineada con Scrum y con el desarrollo ágil de software. (Gómez, 2021).

Herramientas

Para dar soporte al desarrollo de software se han utilizado las siguientes herramientas:

- **Front End:** Se utilizó Javascript, NodeJS, y Bootstrap.
 - **Backend:** Para crear la capa de acceso a datos e interacción con la base de datos no relacional y la blockchain de Filecoin se ha desarrollado un cliente API en lenguaje Javascript, el cual correrá sobre el motor de Node JS.
 - **Smart Contract:** Se ha definido crear un Smart contract en lenguaje Solidity, que tiene por función principal autorizar el acceso a documentos almacenados.
 - **Base de datos:** Para almacenar la información relacionada a los documentos, meta datos e información relevante se utiliza MongoDB. Adicionalmente se decidió utilizar framework de agregadores para desarrollar Pipelines que luego nos permite explotar la información y presentar indicadores claves para Data Analytics
 - **Servidor de desarrollo y Pruebas:** Se utiliza Heroku como servidor de desarrollo y pruebas, por la flexibilidad que tiene, por ser multi plataforma y proveer de manera simple una configuración para utilizar integración continua
- Recolección de datos: A través de observación personal trabajando por más de cuatro años en múltiples implementaciones en San Cristóbal Seguros, luego una experiencia laboral de tres años para la empresa multinacional Guidewire, y finalmente una experiencia internacional de siete meses trabajando para la empresa Business Agility ltd en Reino Unido me han dado una base de información sólida para concluir que los sistemas de gestión de documentos son muy importantes, su integración demanda mucho esfuerzo y se genera un costo importante de mantenimiento que se va incrementando con el correr de los años.
 - Cronograma: La planificación de las actividades se puede observar en la siguiente tabla y diagrama GANTT priorizando cada una de las tareas desarrolladas:

Ilustración 1- Tabla Gantt – Análisis y Documentación

Analisis y Documentación	Estado	Linea de Tiempo	Prioridad	Dependencias
Selección de la temática	Completo	Mar 23 - 31	Alta	-
Introducción	Completo	Apr 1 - 3	Media	Selección de la temática
Justificación	Completo	Apr 4 - 7	Media	Introducción
Objetivos	Completo	Apr 4 - 7	Media	-
Marco teórico referencial	Completo	Apr 7 - 9	Media	Objetivos
Diseño metodológico	Completo	Apr 9 - 15	Media	-
Relevamiento	Completo	Apr 9 - 16	Media	-
Procesos de Negocio	Completo	Apr 16 - 23	Media	-
Diagnostico y Propuesta	Completo	May 3 - 7	Media	-
Objetivos, Límites y Alcance del...	Completo	May 7 - 12	Media	-
Descripción del sistema	Completo	May 12 - 19	Media	Objetivos, Límites y Alcance del prototipo...
Seguridad	Completo	May 24 - 27	Media	Descripción del sistema
Análisis de costos	Completo	May 27 - 29	Media	-
Análisis de riesgos	Completo	May 29 - Jun 1	Media	-
Conclusiones	Completo	Jun 1 - 4	Media	-
Anexos	Completo	Jun 4 - 8	Media	-
Portada, Índice, Referencias	Completo	Jun 14 - 27	Media	-
+ Add		Mar 23 - Jun 27		

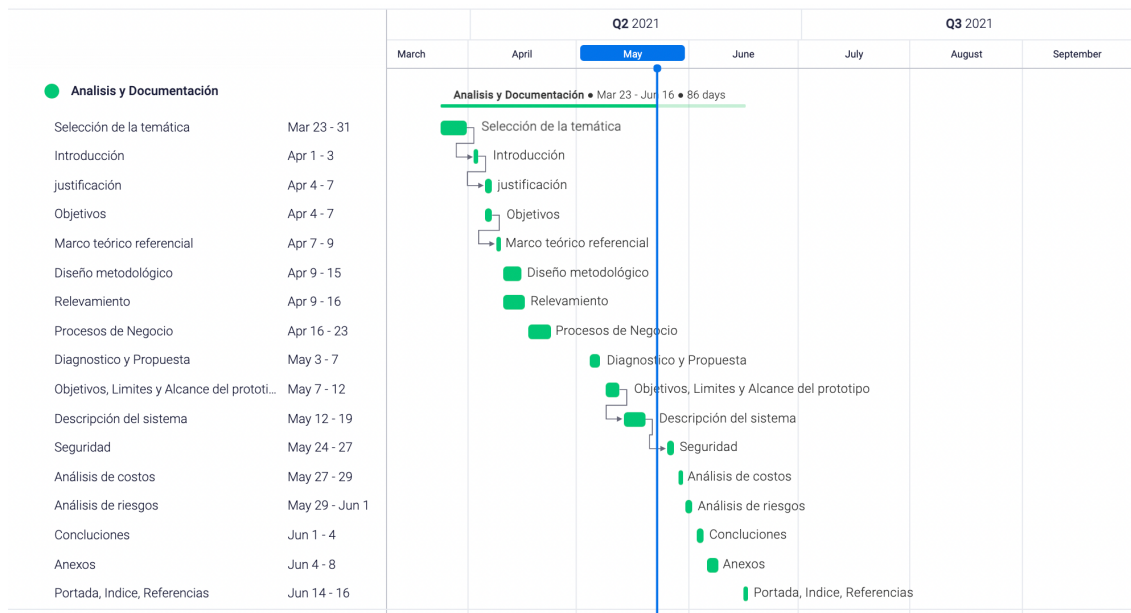
Fuente: Exportado de software Monday.

Ilustración 2 - Tabla Gantt – Desarrollo y Pruebas

Desarrollo y Pruebas	Estado	Linea de Tiempo	Prioridad	Dependencias
Desarrollo de casos de Pruebas	Completo	May 17 - Jun 17	Media	-
Diseño y desarrollo modelo de ...	Completo	May 17 - Jun 11	Media	-
Desarrollo de Backend de Proto...	Completo	May 17 - Jun 17	Media	-
Desarrollo de Front End de Prot...	Completo	May 17 - Jun 19	Media	-
Despliegue en servidor de desa...	Completo	Jun 15 - 26	Media	-
Demo	Completo	Jul 4	Alta	-
+ Add		May 17 - Jul 4		

Fuente: Exportado de software Monday.

Ilustración 3 - Diagrama de Gantt – Análisis y Documentación



Fuente: Exportado de software Monday.

Ilustración 4 - Diagrama de Gantt – Desarrollo y Pruebas



Fuente: Exportado de software Monday.

Relevamiento

El presente relevamiento se realiza sobre una organización modelada, ya que el sistema desarrollado es una solución que puede ser aplicada en múltiples organizaciones de distintos rubros y tamaños, se utiliza como base la estructura de una empresa aseguradora.

Relevamiento Estructural

Las oficinas de la empresa modelada están distribuidas a lo largo de todo el país, teniendo su casa central en Capital Federal y una sucursal en cada provincia.

Cuenta también con una amplia red de productores que interactúan continuamente con los sistemas de gestión de suscripción de pólizas y gestión de siniestros cumpliendo el rol de intermediarios entre el cliente y la empresa.

Debido a esta distribución geográfica y la necesidad de que todas las sucursales estén interconectadas en tiempo real todos sus sistemas están desplegados en una infraestructura basada en la nube.

La organización cuenta con áreas de tecnología, infraestructura, arquitectura de datos, seguridad informática, soporte, y calidad.

Todos los proyectos de TI son administrados por el área de tecnología, los equipos de desarrollo están compuestos por recursos internos de la compañía y por personal contratado a consultoras TI.

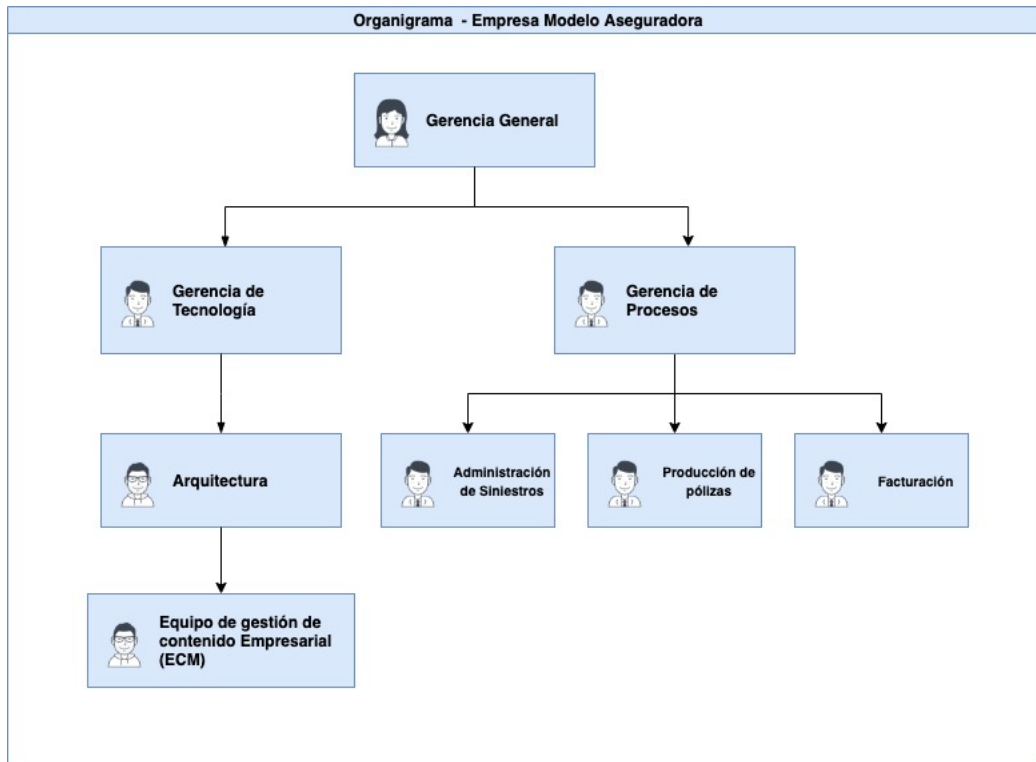
Los sistemas de información implementados fueron desarrollados en parte por personal de la empresa y en parte por equipos subcontratados.

Los procesos que actualmente se encuentran informatizados y tienen soporte de equipos especializados son:

- Gestión de suscripción de Pólizas.
- Gestión de pagos y facturación.
- Administración de contactos.
- Producción de Documentos.
- Gestión de Documentos y Notificaciones.
- Reportes.
- Administración de siniestros.
- Gestión de juicios y litigios.
- Administración de recursos humanos.
- Gestión de recursos humanos.
- Liquidación de sueldos.

Relevamiento Funcional

Ilustración 5 - Organigrama Empresa Modelo



Fuente: Elaboración propia.

Funciones de las áreas

Dentro de cada área definida en el organigrama, podemos definir las siguientes funciones:

- Área Administración de Sinistros: Encargada de gestionar todos los procesos relacionados a la emisión de siniestros, esto abarca la capacidad de recibir notificaciones de los mismos desde clientes y/o entidades intermedias como productores o empresas co-aseguradoras. Almacenar documentación relevante relacionada a los siniestros, y enviar notificaciones al asegurado durante el ciclo de vida del mismo.
- Área Producción de pólizas: Encargada de administrar todos los procesos relacionados a la cotización, contratación de pólizas y las correspondientes

transacciones que puede generar a posteriori un contrato de póliza (Renovaciones, Endosos, Cancelaciones). Que afectan nuestra problemática Podemos mencionar:

- Almacenado y Administración de documentación relevante para el asegurado requeridas por SSN (Súper Intendencia de Seguros de la Nación). Dentro de las cuales se encuentran: constancia de cobertura, caratula de póliza emitida, constancia de endosos, certificado de anulación.
- Área Facturación: Encargada de gestionar los pagos entrantes y salientes de la organización, dentro de las principales funciones que afectan nuestro proyecto podemos mencionar:
 - Almacenamiento y notificación de constancias de pagos.
 - Recibir y procesar información relacionada a pagos realizados por entidades externas que impactan a la organización.
 - Almacenar y proteger información relacionada a medios de pagos de los clientes de la empresa.
- Área Arquitectura: Esta área tiene por objetivo principal de tomar decisiones relacionadas a el diseño de soluciones tecnológicas a nivel general dentro de la empresa, para lograr esto realiza las siguientes funciones:
 - Análisis de impacto en el uso de nuevas tecnologías propuestas desde negocio.
 - Desarrollo de pruebas de concepto y demostración a gerencia al momento de probar alguna nueva solución o cambio que se desea aplicar en los procesos actuales.
 - Encargada de definir los procesos de almacenado, validación y transferencia de información a nivel general, entre aplicaciones ya sean internas o externas.
 - Toma de decisiones con respecto a la tecnología a utilizar para la comunicación entre el sistema Core de la empresa con proveedores externos.
- Gerencia General: Encargada de coordinar todas las áreas principales de la compañía, tomando decisiones de nivel estratégicos. Garantes de la viabilidad económica de los diferentes proyectos llevados a cabo dentro de la empresa.

Encargada de analizar los cambios propuestos desde el sector de tecnología para luego optar por dar curso o no a dichas propuestas.

- Gerencia de Tecnología: Tiene como función principal coordinar los diferentes proyectos tecnológicos dentro de la empresa, sean relacionados al Core principal, o estén relacionados a sistemas satélites necesarios para satisfacer alguna demanda de parte del negocio. Es la que realiza un análisis de factibilidad técnica de las soluciones a presentar.
- Gerencia de Procesos: Tiene como función principal coordinar las áreas relacionadas a los procesos de negocio.
- Equipo de gestión de contenido empresarial (ECM): Equipo de desarrollo, análisis orientado a la interacción entre las áreas de negocio y el proveedor contratado que brinda la solución de gestión de documentación. Dentro de las funciones principales se pueden mencionar:
 - Definir que documentos deben estar afectados por el procesamiento del DMS contratado.
 - Realizar análisis y desarrollo para integrar los sistemas de la empresa a la solución contratada.
 - Definir funciones y actividades que se pueden realizar sobre los documentos almacenados en el DMS.
 - Definir reportes e Indicadores de información que se necesitan explotar para la toma de decisiones relacionadas a la gestión de documentos.

Procesos relevados

Dentro de los procesos afectados por la problemática seleccionamos el caso particular de la constancia de cobertura, aunque aplica para todos los documentos que se generan en los procesos claves del negocio.

Nombre del Proceso: Almacenar Constancia de Cobertura

Roles: Usuario

Pasos: Una vez que el operador del sistema de producción de pólizas finaliza una emisión, se envía información de la misma al sistema de producción de documentos mediante el llamado de un servicio web. Luego de generado el documento, se notifica al sistema DMS (Document Management System) para que almacene el mismo. Para finalizar el proceso,

se registra en el sistema la información relacionada a la metadata del documento almacenado en el DMS

Nombre del Proceso: Recuperar Constancia de Cobertura

Roles: Usuario

Pasos: El operador del sistema de producción de pólizas desea recuperar el documento constancia de cobertura almacenado en el dms, para enviar una copia digital al cliente. Mediante la llamada a un servicio web, a demanda, se recupera el documento almacenado en el DMS, enviando como request el id de la transacción, un token que, garantiza la autorización del usuario para acceder al documento, y se procede a descargar el documento.

Nombre del Proceso: Visualizar reporte de documentos almacenados en el DMS por Usuario

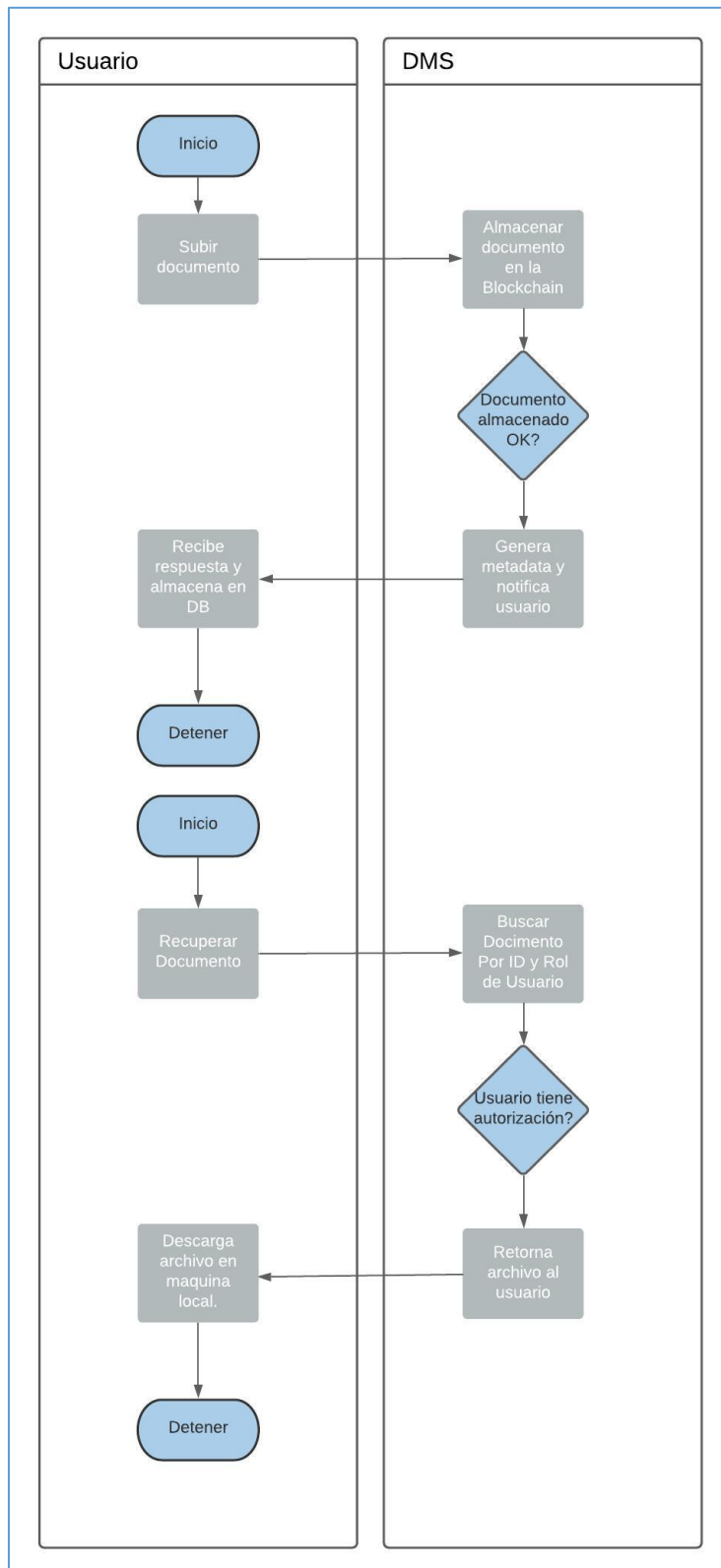
Roles: DMS

Pasos: Como administrador del sistema de administración de documentos, se solicita ver información estadística de volumen de documentos almacenados por usuarios para utilizar como fuente de datos en reporte solicitado por Gerencia de procesos.

Procesos de Negocio

Basándonos en el relevamiento funcional se desarrolla un flujo de proceso genérico en el cual se pueden observar las funcionalidades de almacenamiento y recuperación de un documento.

Ilustración 6 - Diagrama de flujo - Almacenar y Recuperar Documento



Fuente: Elaboración propia.

Diagnóstico y Propuesta

Diagnóstico

Tabla 1- Proceso Almacenar constancia de cobertura

Nombre del proceso: Almacenar constancia de cobertura	
Problemas	Causas
<p>Por la naturaleza de los sistemas de la organización, todo el almacenamiento esta gestionado en la nube, lo que con el tiempo termina generando un costo importante de mantenimiento debido al incremento periódico de la documentación almacenada y la necesidad de preservar la misma por largos periodos de tiempo.</p>	<p>Al ser un proceso relacionado a una transacción masiva, que no solo afecta a la emisión de una póliza, sino también a anulación, endosos y re-factoraciones Se termina generando un gran volumen de documentos generados.</p> <p>También se pueden encontrar casos particulares de pólizas con miles de certificados, lo que incrementa el volumen aun más.</p>

Fuente: Elaboración propia.

Tabla 2- Proceso recuperar constancia de cobertura

Nombre del proceso: Recuperar constancia de cobertura	
Problemas	Causas
Debido a la característica del almacenado centralizado, se corre el riesgo que en caso de verse afectados los servidores en donde se guarda la información de los documentos, sea por un ataque informático o eventos externos, se pierda acceso parcial o total a la información. Si las credenciales del usuario se ven comprometidas, se podría perder la confidencialidad del documento.	A demanda, los usuarios y/o intermediarios solicitan descargar o visualizar la constancia de cobertura correspondiente a un contrato de póliza, endoso o renovación. Dicha información es confidencial y se debe garantizar que solo los usuarios con los permisos adecuados tengan acceso al mismo. Actualmente los documentos se almacenan Base de datos centralizadas. Con respecto a la seguridad, solo se garantiza el acceso al documento validando los permisos que el usuario tiene a la interfaz gráfica del sistema.

Fuente: Elaboración propia.

Tabla 3 - Proceso visualizar reporte de documentos almacenados

Nombre del proceso: Visualizar reporte de documentos almacenados en el DMS por usuario	
Problemas	Causas
Cometer errores en la toma de decisiones sobre procesos que afectan la gestión de documentos y se pierde la posibilidad de optimizar costos y recursos.	El área de tecnología no cuenta con información precisa sobre los volúmenes de almacenamiento generados por los procesos de generación y administración de documentos. Tampoco tienen información acerca de que tipo de documentos son los que generan el mayor volumen de documentos.

Fuente: Elaboración propia.

Propuesta

Debido a que la organización modelada en el relevamiento contaba con una estructura compuesta por múltiples oficinas ubicadas físicamente en distintas ciudades, y teniendo en cuenta que gran parte de los empleados trabajan de forma remota, se ha desarrollado un prototipo de aplicación web con almacenamiento distribuido, el cual a su vez contiene una capa de servicios que facilita la integración con cualquier sistema externo que desee comunicarse y consumir las funcionalidades expuestas.

Para lograr un bajo costo de almacenamiento se ha utilizado la blockchain del proyecto Filecoin como base de datos distribuida, lo que adicionalmente nos brinda una mejora sustancial en la integridad y confidencialidad de la información.

Para poder interactuar con la aplicación se ha diseñado una serie de interfaces que permiten cubrir los requerimientos de registro de usuarios al sistema, gestionar los permisos de los mismos, realizar operaciones de almacenado y consulta de documentos.

Con el objetivo de garantizar la confidencialidad y acceso a la documentación se desarrolló un Smart contract que tiene la responsabilidad de gestionar la autorización de un usuario a la documentación asociando las credenciales del usuario presentes en la metadata del documento con una dirección de ethereum.

Se ha definido una interfaz de administración que permite generar reportes con información estadística relevante para la toma de decisiones de la gerencia.

Objetivo, Límite y Alcances del Prototipo

Objetivo del Prototipo

Administrar de manera segura e íntegra documentación digital sensible y brindar reportes con información relevante para la toma de decisiones estratégicas tanto en organizaciones públicas como privadas.

Límites

El sistema abarca desde que un usuario se registra al sistema, y posteriormente estando autenticado decide subir un documento electrónico al sistema, garantizando el almacenamiento y administración del mismo de manera segura, resguardando la integridad y confidencialidad de la información durante todo el proceso.

Alcances

Teniendo en cuenta que la solución a desarrollar puede aplicar a diversos actores se define una lista de procesos genéricos que en su conjunto pueden brindar solución a múltiples problemáticas.

Los procesos definidos son:

- Registro y autenticación de Usuarios.
- Desarrollo de Smart contract para autorizar accesos a recursos.
- Gestión de roles y permisos.
- Almacenamiento de Documentos de manera descentralizada.
- Gestión de Documentos almacenados.
- Generación de reporte con información estadística del volumen de documentos almacenados por usuarios.
- Generación de reporte con información estadística del volumen de documentos almacenados por tipo de documento.

Descripción del sistema

Product Backlog

Ilustración 7 - Product Backlog

ID	Estado	Puntos	Impacto	Nombre
TFG-001	Completo	5 Puntos	Medio	Registro de Usuario
TFG-002	Completo	8 Puntos	Medio	Autenticación de Usuario
TFG-003	Completo	13 Puntos	Alto	Upload de Documentos
TFG-004	Completo	8 Puntos	Alto	Listar Documentos
TFG-005	Completo	5 Puntos	Medio	Descarga de Documentos
TFG-006	En progreso	8 Puntos	Medio	Generar Reportes
TFG-007	En progreso	3 Puntos	Medio	Infraestructura
TFG-008	En progreso	3 Puntos	Medio	Infraestructura
TFG-009	Completo	5 Puntos	Alto	Autenticación de Usuario
TFG-010	Completo	5 Puntos	Alto	Gestión de Documentos

Fuente: Extraído de software Monday.

Historias de Usuario

Tabla 4 - Registro de Usuario

ID	TFG-001	Nombre	Registro de Usuario
Descripción		Como usuario web quiero registrarme al sistema para poder gestionar documentos digitales.	
Criterios de aceptación		<p>Dada una petición de registro en el formulario de alta de usuario.</p> <p>Cuando el usuario haya completado todos los campos requeridos, los datos sean validos, y el Smart contract haya vinculado los datos del usuario con la dirección auto generada en la red de ethereum.</p> <p>Entonces se mostrará un mensaje que indique que el registro se realizó con éxito.</p>	
Prioridad	Media	Puntos de historia	5

Fuente: Elaboración propia.

Tabla 5 - Login de Usuario

ID	TFG-002	Nombre	Login de Usuario
Descripción		Como usuario registrado del sistema quiero autenticarme para poder subir documentos al sistema.	
Criterios de aceptación		<p>Dada una petición de autenticación al sistema en la pagina de login.</p> <p>Cuando los datos ingresados sean validos y el Smart contract que corre en la red de ethereum garantice el acceso.</p> <p>Entonces se mostrará la pagina principal del sistema con las funcionalidades habilitadas para el usuario.</p>	
Prioridad	Media	Puntos de historia	8

Fuente: Elaboración propia.

Tabla 6 - Upload de Documento

ID	TFG-003	Nombre	Upload de Documento
Descripción		Como usuario autenticado al sistema quiero subir un documento.	
Criterios de aceptación		<p>Dada una petición de upload de documento desde la pantalla de gestión de documentos del usuario.</p> <p>Cuando el documento cumpla con los requisitos de limites de tamaño, formato y la petición al servicio web de almacenamiento retorne una respuesta de éxito.</p> <p>Entonces se cargará la pagina de gestión de documentos y se podrá ver en el listado de documentos el documento recientemente almacenado.</p>	
Prioridad	Media	Puntos de historia	13

Fuente: Elaboración propia.

Tabla 7 - Listar Documentos

ID	TFG-004	Nombre	Listar Documentos
Descripción		Como usuario autenticado al sistema quiero ver la lista de documentos subidos al sistema.	
Criterios de aceptación		Dada una petición a la pagina de documentos de usuario. Cuando el usuario haya subido previamente documentos. Entonces se mostrara una lista de documentos almacenados ordenados por fecha de creación en orden descendiente.	
Prioridad	Media	Puntos de historia	5

Fuente: Elaboración propia.

Tabla 8 - Descarga de Documento

ID	TFG-005	Nombre	Descarga de Documento
Descripción		Como usuario autenticado en el sistema quiero descargar un documento.	
Criterios de aceptación		Dada una petición a de descarga de documento. Cuando el Smart contract garantice que el usuario tenga el correcto nivel de acceso de descarga. Entonces se iniciará la descarga del archivo en una nueva pestaña del navegador.	
Prioridad	Media	Puntos de historia	5

Fuente: Elaboración propia.

Tabla 9 - Generar Reporte

ID	TFG-006	Nombre	Generar Reporte
Descripción	Como usuario administrador autenticado en el sistema quiero visualizar reporte de documentos almacenados por usuario.		
Criterios de aceptación	<p>Dada una petición a de generación de reportes de documentos por usuario.</p> <p>Cuando el Smart contract garantice que el usuario tenga el correcto nivel de acceso para la generación del reporte.</p> <p>Entonces se visualizará un grafico estadístico con el reporte solicitado en la pantalla de Documentos almacenados por usuario.</p>		
Prioridad	Media	Puntos de historia	5

Fuente: Elaboración propia.

Sprint Backlog

Ilustración 8 - Sprint Backlog - parte 1

Nombre	Subitems	REQ-ID	Puntos	Impacto	Nombre
Como usuario deseo registrarme al sistema	Desarrollo Interfaz de Usuario, Definir modelo de datos, Desarrollar servicio y controlador para almacenar la información	TFG-001	5	Medio	Registro de Usuario
Subitems	Name	Owner	Puntos	TASK ID	Prioridad
	Desarrollo Interfaz de Usuario	Mario Andrés Bedin	1	TASK-001	medio
	Definir modelo de datos	Mario Andrés Bedin	1	TASK-002	medio
	Desarrollar servicio y controlador para almacenar la información	Mario Andrés Bedin	3	TASK-003	medio
Como usuario registrado deseo loguearme al sistema	Desarrollar interfaz de usuario, Definir modelo de daos, Desarrollar servicio y controlador para autenticar el usuario., Integrar servicio de Auth0 para generar token de autorización	TFG-002	8	Medio	Autenticación de Usuario
Subitems	Name	Owner	Puntos	TASK ID	Prioridad
	Desarrollar interfaz de usuario	Mario Andrés Bedin	1	TASK-004	medio
	Definir modelo de daos	Mario Andrés Bedin	1	TASK-005	medio
	Desarrollar servicio y controlador para autenticar el usuario.	Mario Andrés Bedin	2	TASK-005	medio
	Integrar servicio de Auth0 para generar token de autorización	Mario Andrés Bedin	3	TASK-006	alta

Fuente: Elaboración propia.

Ilustración 9 - Sprint Backlog - parte 2

Como usuario autenticado deseo subir un documento al sistema	Desarrollar Interfaz de usuario, Definir metadata del documento, Desarrollar servicio para almacenar documento en la blockchain de Filecoin, Implementar integración con smart contract para autorizar usuario	TFG-003	13	Alto	Upload de Documentos
Subitems	Name	Owner	Puntos	TASK ID	Prioridad
	Desarrollar Interfaz de usuario	Mario Andrés Bedin	1	TASK-007	medio
	Definir metadata del documento	Mario Andrés Bedin	2	TASK-008	medio
	Desarrollar servicio para almacenar documento en la blockchain de Filecoin	Mario Andrés Bedin	5	TASK-009	alta
	Implementar integración con smart contract para autorizar usuario	Mario Andrés Bedin	5	TASK-010	alta
Como usuario autenticado deseo ver la lista de documentos subidos al sistema	Desarrollar servicio que recupere y liste documentos por usuario, Implementar integración con smart contract para autorizar usuario	TFG-004	8	Alto	Listar Documentos
Subitems	Name	Owner	Puntos	TASK ID	Prioridad
	Desarrollar servicio que recupere y liste documentos por usuario	Mario Andrés Bedin	3	TASK-011	medio
	Implementar integración con smart contract para autorizar usuario	Mario Andrés Bedin	5	TASK-012	alta

Fuente: Elaboración propia.

Ilustración 10 - Sprint Backlog - parte 3

Como usuario autenticado deseo descargar documento del sistema	Desarrollar servicio que permita descargar archivo al directorio local del usuario, Implementar integración con smart contract para autorizar descarga al usuario	TFG-005	5	Medio	Descarga de Documentos
Subitems	Name	Owner	Puntos	ID	Prioridad
	Desarrollar servicio que permita descargar archivo al directorio local del usuario	Mario Andrés Bedin	2	TASK-013	medio
	Implementar integración con smart contract para autorizar descarga al usuario	Mario Andrés Bedin	3	TASK-014	alta

Fuente: Elaboración propia.

Ilustración 11 - Sprint Backlog - parte 4

Como usuario administrador necesito visualizar reporte de cantidad de documentos almacenados por usuario	Desarrollar interfaz de usuario administrador, Modelar reporte consolidando información de metadata y usuarios, Desarrollar servicio que se conecte con mongoDB y permita previsualizar el reporte solicitado	TFG-006	8	Medio	Generar Reportes
Subitems	Name	Owner	Puntos	ID	Prioridad
	Desarrollar interfaz de usuario administrador	Mario Andrés Bedin	1	TASK-015	
	Modelar reporte consolidando información de metadata y usuarios	Mario Andrés Bedin	2	TASK-016	
	Desarrollar servicio que se conecte con mongoDB y permita previsualizar el reporte solicitado	Mario Andrés Bedin	5	TASK-017	

Fuente: Elaboración propia.

Ilustración 12 - Sprint Backlog - parte 5

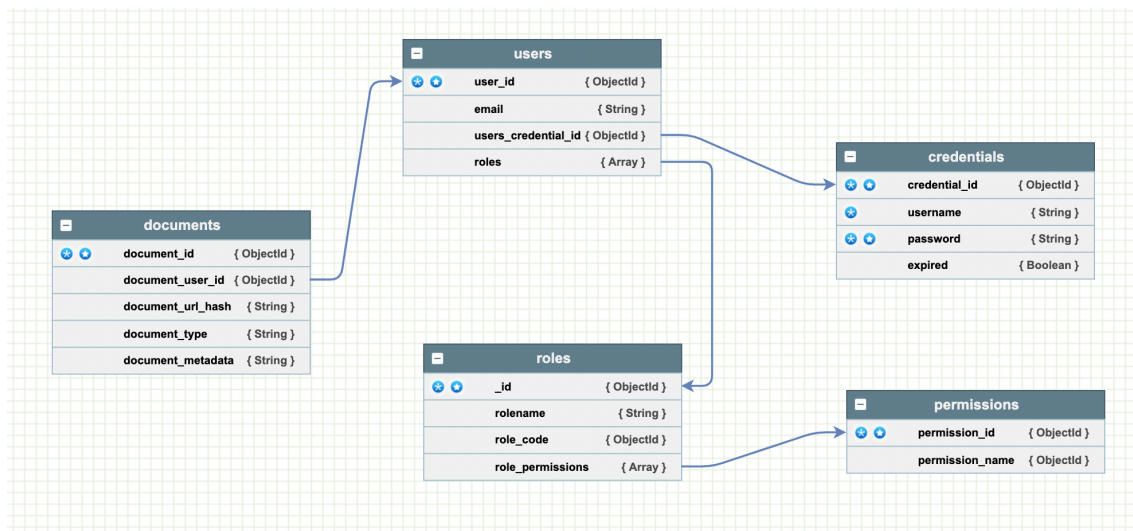
Configurar Instancia EC2 en Amazon para poder desplegar una version del sistema	TFG-007	3	Medio	Infraestructura
Configurar Instancia MongoDB cloud para poder almacenar las credenciales y la metadata de los documentos almacenados en el sistema	TFG-008	3	Medio	Infraestructura
Desarrollar Smart Contract que garantice los permisos de un usuario para el acceso a los documentos del sistema	TFG-009	5	Alto	Autenticación de Usuario
Configurar nodo Filecoin en modo desarrollo para permitir el almacenado descentralizado de documentos	TFG-010	5	Alto	Gestión de Documentos
		109		

Fuente: Elaboración propia.

Estructuras de Datos

Para almacenar la información del prototipo vamos a utilizar una base de datos no relacional sobre el cliente MongoDB con la siguiente estructura.

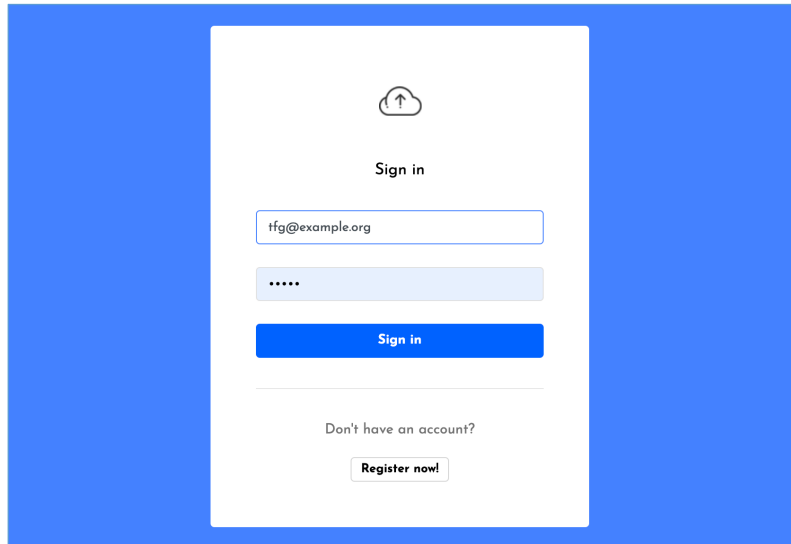
Ilustración 13 - Estructura de Datos



Fuente: Elaboración propia.

Prototipos de Interfaces de Pantallas

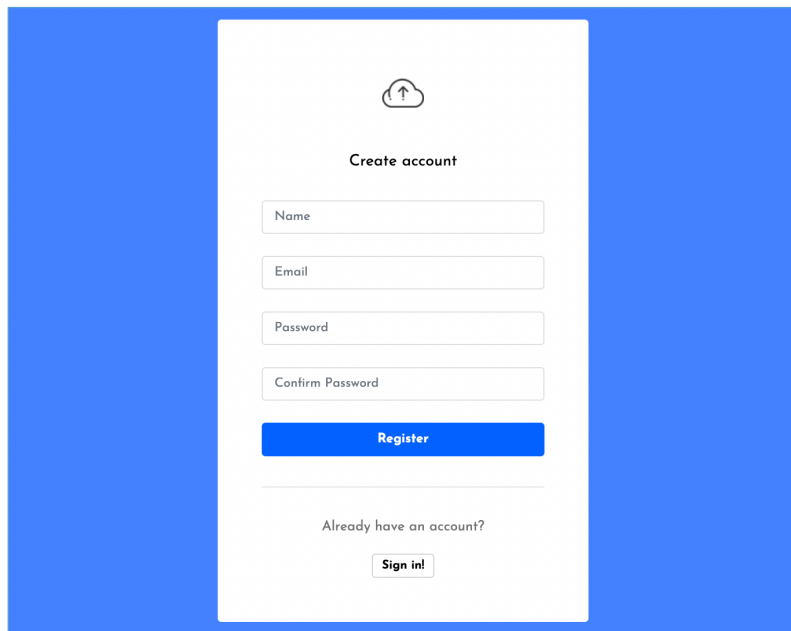
Ilustración 14 - Pantalla Login



A login screen prototype with a white background centered on a blue background. At the top is a cloud icon with an upward arrow. Below it is the text "Sign in". There are two input fields: the first contains the email "tfg@example.org" and the second contains six dots. A blue "Sign in" button is below the password field. A horizontal line separates the login section from the registration section. Below the line is the text "Don't have an account?" and a "Register now!" button.

Fuente: Elaboración propia

Ilustración 15 - Pantalla Registro de Usuario



A user registration screen prototype with a white background centered on a blue background. At the top is a cloud icon with an upward arrow. Below it is the text "Create account". There are four input fields: "Name", "Email", "Password", and "Confirm Password". A blue "Register" button is below the "Confirm Password" field. A horizontal line separates the registration section from the login section. Below the line is the text "Already have an account?" and a "Sign in!" button.

Fuente: Elaboración propia

Ilustración 16 - Pantalla Upload Documento

The screenshot shows a 'New Document' form with the following elements:

- Title:** A text input field.
- Description:** A larger text area with a small icon for text formatting.
- File Selection:** A button labeled 'Choose File' next to the text 'No file chosen'.
- Actions:** Two buttons at the bottom: 'Save' (blue) and 'Back' (red).

Fuente: Elaboración propia

Ilustración 17 - Pantalla Listar Documentos

The screenshot displays a 'Files' list with the following data:

NAME	UPLOADED	CID
Testing 1 350 KB	30/10/2021	bafybeick3m5u3kooatn-fak5m62z4f3gche3yripfzdvkvw
Loaded from sample data1 350 KB	29/10/2021	QmXbKQJYBuNUnsqF7YVWuSok492CH4HqMqKLYU
Loaded from sample data1 350 KB	29/10/2021	QmaAJ-OpQ5MCCYksw/YeR9JzWZdKsXUE3M4Utl4P
Loaded from sample data1 350 KB	29/10/2021	QmRHRGkaQkXfniU0V9yEzZ5VY2u2buNR6dP3-yjy
Loaded from sample data1 350 KB	29/10/2021	QmGZYbp-vYNgglBubAL6YU8MXXqkMukCOFjL3Dz
test ipfs 350 KB	31/07/2021	bafybeickumord3gq4f5wofsq7m44z6m4d5m6k47j2bnk4u4

Storage Overview:

- 35% Storage Usage
- 35% Storage Usage
- Images: 15.7 GB (259 Files)
- Videos: 20 GB (8 Files)
- Documents: 10.5 GB (46 Files)
- Other Files: 2.8 GB (50 Files)

Fuente: Elaboración propia

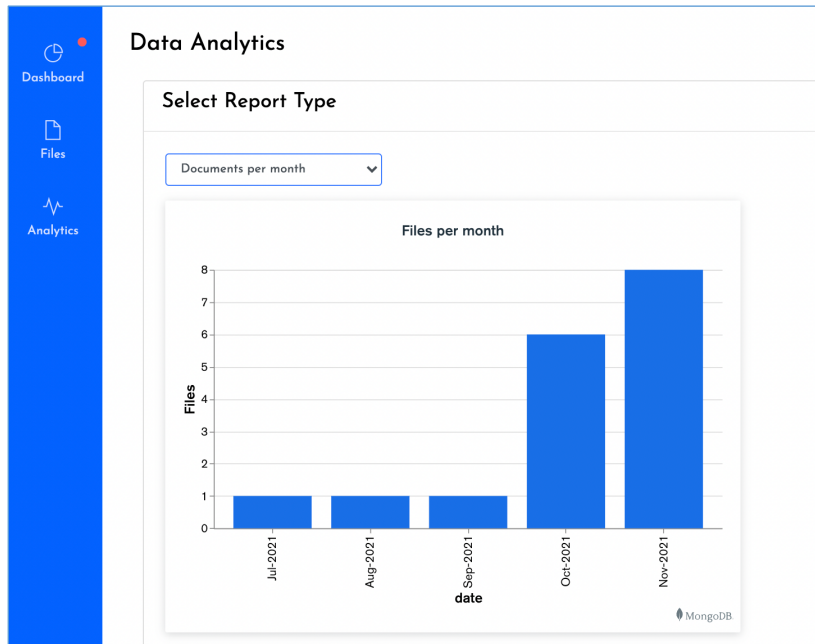
Ilustración 18 - Lista de archivos en Nodo IPFS (Filecoin Hot Layer)

The screenshot shows the IPFS interface with the following file listing:

File Name	CID	Size
Index of /ipfs/bafybeickphmz3tmpvumroqnxzfyn7daza5pni6oahns5zh1me55vmsui		12 MB
..		
.keep	QmUN...A3Nn	4 B
IPFS Document	bafy...7zwl	513 kB
Testing 1	bafy...kanu	1.6 MB
defl-beginer.pdf	bafy...vay4	9.3 MB
test	QmUk...4KQ2	68 B
test.pdf	bafy...ltwe	724 kB

Fuente: Fleek Co. Proveedor de hosting vps.

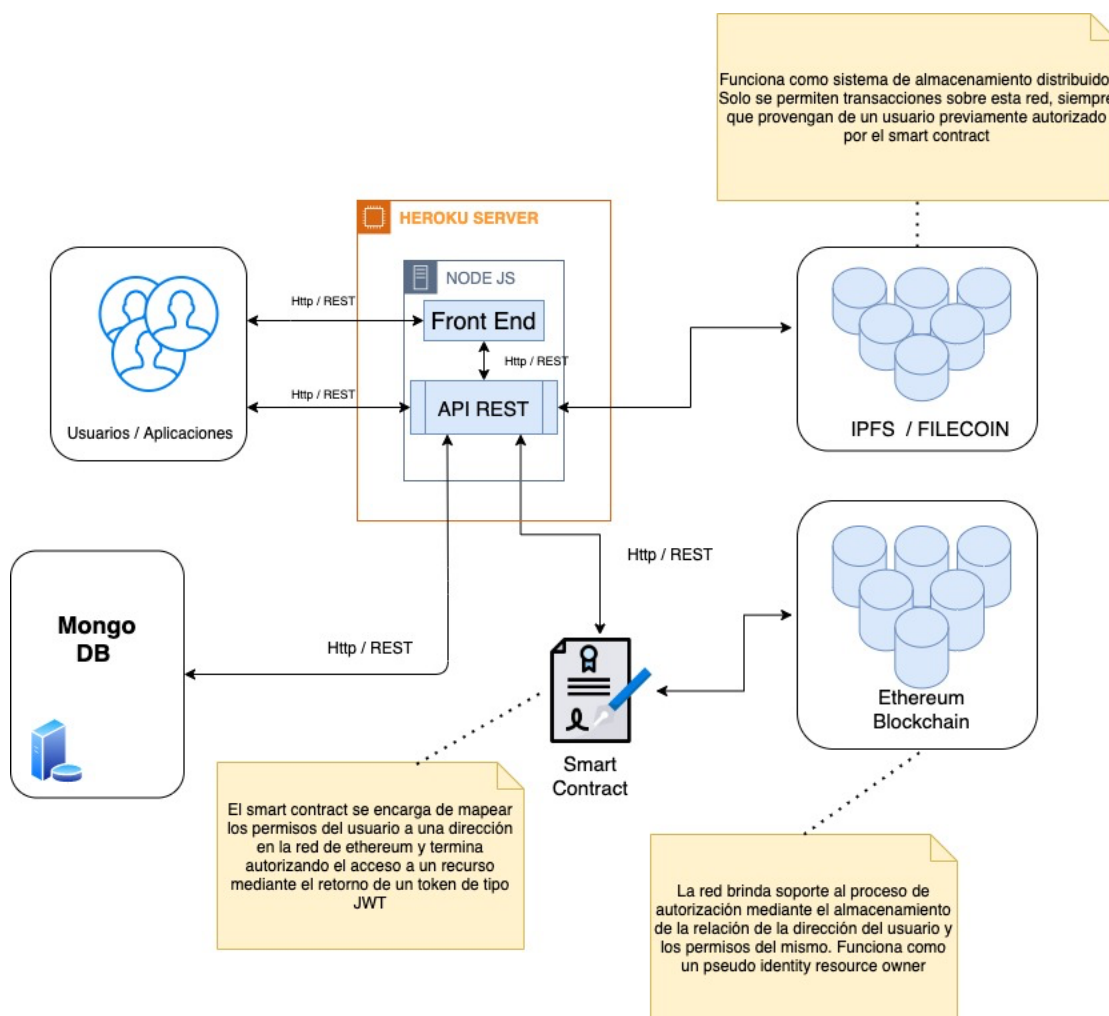
Ilustración 19 - Reporte generado mediante analítica de datos.



Fuente: Elaboración propia.

Diagrama de Arquitectura

Ilustración 20 - Diagrama de Arquitectura



Fuente: Elaboración propia

Seguridad

Acceso a la Aplicación

En cuanto a las políticas de seguridad de acceso lógico a la aplicación, con el objetivo de potenciar la seguridad y confidencialidad de los datos intercambiados durante el acceso al sistema, se han definido dos capas de seguridad, que se detallan a continuación:

- Capa de autenticación: En esta capa se ha determinado utilizar el protocolo OIDC 2.0 el cual es una extensión del protocolo OAuth2 (Sitio oficial de Auth0, 2021) con el cual se asegura la autenticación mediante un mecanismo basado en

encriptación asimétrica, con la cual por medio de una clave pública y un proceso de encriptación se obtiene un token utilizando el estándar JWT (Sitio oficial de JWT, 2021) que garantiza el acceso a la aplicación.

- Capa de autorización: Para el proceso de autorización, se han definido roles y privilegios asignables a cada usuario, con el objeto de restringir determinadas secciones del sistema que solo deberían estar disponibles a usuarios especiales. Este mecanismo es garantizado por medio de un Smart Contract que corre sobre la blockchain Ethereum y es enlazado a la capa de aplicación por medio de un IDP 'Identity Provider' (Sitio oficial de Okta, 2021).

Con respecto a los datos requeridos como input del proceso de autenticación y posterior autorización se han establecido los siguientes:

- Usuario y password: Para el dato usuario, se exige que el mismo sea una cuenta de email y que el formato ingresado sea válido. Con respecto al dato password, el criterio de validación es que el mismo tenga al menos ocho caracteres, al menos una letra mayúscula y al menos un carácter especial.
- Expiración de password: Se define una política de expiración de password de cuatro meses.
- Asignación de Roles a Usuarios: A cada usuario que se registra al sistema se le asigna uno o más roles para determinar el grado de acceso que tiene sobre la aplicación.
- Asignación de dirección sobre Ethereum: El sistema genera de forma automática una dirección en la red de Ethereum que se asocia a los datos del usuario registrado con el objetivo de desacoplar el registro básico del usuario al proceso de autenticación que gestiona el Smart Contract.

En relación a la seguridad de acceso a base de datos se ha definido lo siguiente:

- Credenciales: Bajo ningún aspecto se almacena información relacionada a usuario, contraseña y url de conexión relacionadas a la base de datos en texto plano. Dicha información se configura mediante variables de ambiente en el servidor de aplicaciones y se visualizan de manera encriptadas en el mismo.
- Roles: Se definen distintos roles a ser configurados en los usuarios que acceden a la base de datos, con el objetivo de restringir funcionalidades sobre la base de datos.

Políticas de respaldo

Como estrategias a seguir en caso de catástrofes o incidentes relacionados a pérdida de información por errores humanos se ha definido las siguientes políticas:

- Política de aislamiento: Los procesos que generan los backups de la base de datos deben estar totalmente desacoplados del sistema desplegado en el ambiente productivo.
- Disminución de impacto: Para disminuir el impacto que puede producir el proceso de backups sobre la aplicación, se define programar la ejecución del mismo fuera de los horarios pico de uso del sistema. Con esto se logra mitigar una posible degradación de la performance en el uso de la aplicación.
- Proceso de recuperación: La estrategia de backup es buena solo si el mecanismo para restaurar la información comprometida se encuentra correctamente definido, por esto se han establecido dos políticas de recuperación:
 - Recuperación automática: Se ha definido un evento disparador del proceso llamado ‘data emergency’ que es disparado en caso de catástrofes.
 - Recuperación a demanda: Mediante la funcionalidad ‘mongorestore’ nativa del cliente utilizado para el almacenamiento de datos.
- Tiempos de recuperación: Se estableció como parámetro de tiempo deseable, que el sistema pueda estar operativo transcurrido al menos una hora desde el episodio de pérdida de información. Este parámetro es clave en la estrategia y también se lo puede definir como RTO (Recovery time objective).
- Punto de recuperación: Se definió el parámetro de máxima cantidad de información dispuesta a perder (Expresada en tiempo), también llamado RPO (Recovery point objective), con el valor de 24 horas. Esto significa que se van a realizar backup del sistema diariamente.

Análisis de Costos

A continuación, se detallan los costos necesarios para llevar adelante la implementación del proyecto, estos costos están organizado en las categorías: Hardware, Software y Recursos Humanos.

Los montos en dólares se convierten a pesos utilizando el tipo de cambio publicado por el Banco Central de la Republica Argentina el día 11 de junio del año 2021.

Hardware

La información del costo de cada ítem de la siguiente tabla, fue extraída de las siguientes fuentes:

- AWS pricing calculator (Sitio oficial de AWS, 2021).
- Auth0 sitio oficial (Sitio oficial Auth0, 2021).
- Lenovo Argentina sitio oficial (Sitio oficial Lenovo Argentina, 2021).

Tabla 10 - Costos de Hardware

Ítem	Detalle	Cantidad	Precio Unitario	Total
Cloud Server (Producción)	Instancia Amazon EC2 con Sistema Operativo Linux. 8 vCPU, 32GB Ram. Modo auto escalado.	1	30 USD (Mensual)	4950 ARS
Cloud Server (Testing)	Instancia Amazon EC2 con Sistema Operativo Linux. 6 vCPU, 16GB Ram. Modo auto escalado.	1	15 USD (Mensual)	2470 ARS
Cloud Server (Desarrollo)	Instancia Amazon EC2 con Sistema Operativo Linux. 4 vCPU, 16GB Ram. Modo auto escalado.	1	10 USD (Mensual)	1650 ARS
IDP Server (Identity Provider)	Plan B2C Auth0	1	23 USD (Mensual)	3.795 ARS
Laptop para desarrollo.	Lenovo Thinkpad X1 Carbon. CPU Intel Core i7, 32GB DDR4, Disco SSD 1 TB.	3	313.999 ARS	941.997 ARS

Laptop para testing.	Lenovo Thinkpad E15. CPU Intel Core i7, 32GB DDR4, Disco SSD 512 GB.	2	171.999 ARS	343.998 ARS
Total				1.298.860 ARS

Fuente: Elaboración propia

1 USD = 100 ARS (A esta cotización se incrementa un 30% de impuesto PAIS y 35% por recargo de ganancias)

Software

La información del costo de cada ítem de la siguiente tabla, fue extraída de las siguientes fuentes:

- Sitio oficial Microsoft Argentina (<https://www.microsoft.com>).
- Sitio oficial JetBrains (<https://www.jetbrains.com/idea>).
- Sitio oficial Monday (<https://monday.com/pricing>)

Tabla 11 - Costos de Software

Detalle	Cantidad	Precio Unitario	Total
Licencia Windows 10 Pro (Licencia empresarial)	4	21.999 ARS	87.996 ARS
Licencia IntelliJ IDEA studio	3	14,90 USD	2.458,5 ARS
Plan Standard de software de gestion de Proyecto Monday	1	10 USD	1.650 ARS
Total			92.104,5 ARS

Fuente: Elaboración propia

1 USD = 100 ARS (A esta cotización se incrementa un 30% de impuesto PAIS y 35% por recargo de ganancias)

Recursos Humanos

De acuerdo al cronograma planificado, a continuación, se detallan los costos de los perfiles necesarios para la implementación del sistema. La información del costo de la hora de cada uno de los perfiles se obtuvo del sitio oficial del CPCIPC - Consejo Profesional de ciencias informáticas de Córdoba (<https://www.cpcipc.org.ar/content/honorarios>).

Tabla 12 - Costos de Recursos Humanos

Perfil	Cantidad	Horas	Valor Hora	Total
Analista de Negocio	1	480	4.265 ARS	2.047.200 ARS
Tester	1	120	3.129 ARS	375.480 ARS
Desarrollador Backend	1	400	5.152 ARS	2.060.800 ARS
Desarrollador Frontend	1	400	3.490 ARS	1.396.000 ARS
Sys Admin	1	120	4.709 ARS	565.080 ARS
Total				6.444.560 ARS

Fuente: Elaboración propia

Tabla 13 - Costos Totales

Categoría de Costo	Precio
Hardware	1.298.860 ARS
Software	92.104,5 ARS
Recursos Humanos	6.444.560 ARS
Costo Total	7.835.524,5 ARS

Fuente: Elaboración propia

Análisis de Riesgos

Para detectar la mayor cantidad de escenarios en los que se debería aplicar acciones de contingencia se presenta a continuación una tabla en donde se clasifican los riesgos que pueden afectar el cumplimiento de los objetivos.

Tabla de riesgos

Tabla 14 - Listado de Riesgos.

Código	Riesgo	Detalle
Risk-001	Requerimientos de mala calidad	Las definiciones funcionales incompletas o con errores de interpretación pueden generar re trabajos indeseados durante la etapa de desarrollo y bloqueos de avance importantes.
Risk-002	Desarrolladores con bajo seniority	Si los recursos no cuentan con la deseada experiencia y capacidad, se pueden generar demoras significantes en el avance del desarrollo, así como también problemas de performance, y mayor cantidad de errores reportados en las etapas de testing e implementación.
Risk-003	Estrategia de Testing insuficiente	Si la estrategia de testing no es la correcta, y a su vez no se realizan pruebas suficientes, se puede pasar por alto errores que generen re trabajo posteriormente o incluso errores que se encuentren cuando la aplicación ya se encuentre en producción, lo que nos obligaría a tratarlos como problemas críticos.

Risk-004	Disponibilidad de servidores interrumpida.	A pesar de contratar servicios de alta disponibilidad para desplegar aplicación y base de datos. Se pueden generar interrupciones que afecten el normal uso de la aplicación.
Risk-005	Configuración de seguridad deficiente en los servidores.	Si la configuración de seguridad no se realiza correctamente y siguiendo las mejores practicas, los servidores pueden quedar vulnerables a ataques externos.
Risk-006	Estrategia de backup incorrecta.	Si no se define una estrategia de generación de copias de respaldo adecuada, pueden generarse grandes perdidas de información en situaciones de catástrofe.
Risk-007	Utilización de tecnología en una etapa de evolución y madurez muy temprana.	Al utilizar tecnología muy nueva (Blockchain) el desarrollo puede verse afectado por la necesidad de invertir mucho tiempo en investigación sobre situaciones o bloqueos no contemplados.
Risk-008	Posibles cambios normativo.	El ámbito normativo sobre la tecnología utilizada esta actualmente siendo analizado y cualquier cambio de alto impacto sobre la misma, podría generar demoras, re trabajos e incluso cambios de gran escala en la solución planteada.

Fuente: Elaboración propia

Clasificación de riesgos

Para la clasificación y posterior gestión de los riesgos previamente detallados, se utilizó el sistema FMEA 'Failure Mode and Effective Analysis' (Sitio oficial de la organización America Society for Quality, 2021) en donde se clasifican por los siguientes criterios:

- Frecuencia: se establece puntuación de 1 a 10 (Menos frecuente a mayor cantidad de ocurrencias respectivamente).
- Gravedad: se establece puntuación de 1 a 10 (Menor gravedad a mayor gravedad respectivamente).
- Detección: se establece puntuación de 1 a 10 (Detectable de manera simple a mayor dificultad para ser detectado respectivamente).
- RPN (Risk Priority Number): resultado de multiplicar frecuencia x gravedad x detección

Tabla 15 - Clasificación de Riesgos

Código	Riesgo	Gravedad	Frecuencia	Detección	RPN
Risk-001	Requerimientos de mala calidad	8	3	7	168
Risk-002	Desarrolladores con bajo seniority	7	4	5	140
Risk-003	Estrategia de Testing insuficiente	7	5	4	140
Risk-004	Disponibilidad de servidores interrumpida.	9	2	1	18
Risk-005	Configuración de seguridad deficiente en los servidores.	8	1	5	40

Risk-006	Estrategia de backup incorrecta.	8	1	7	56
Risk-007	Utilización de tecnología en una etapa de evolución y madurez muy temprana.	5	2	8	80
Risk-008	Posibles cambios normativo.	9	1	6	54

Fuente: Elaboración propia

Luego de realizar la priorización de cada uno de los riesgos encontrados, se pretende con las siguientes acciones de contingencia reducir notablemente la probabilidad de ocurrencias de los riesgos mencionados de manera integral.

Dichas acciones se ordenan teniendo en cuenta la prioridad del riesgo que intentan mitigar:

- Risk-001: Definir como parte del proceso de definición de requisitos las tareas de verificación y validación de los mismos para asegurar un grado aceptable de calidad
- Risk-002: Contratar capacitaciones relacionadas a la tecnología utilizada durante el desarrollo.
- Risk-003: Definir un manual de buenas prácticas con respecto a las estrategias de Testing a aplicar. Adicionalmente comenzar a realizar pruebas de manera temprana.
- Risk-007: Realizar workshop en donde se presentan las ventajas y desventajas del uso de la tecnología en cuestión, junto con un detalle de las potenciales utilidades de la misma, y casos de uso exitosos en implementaciones similares.

- Risk-006: Definir una estrategia de backups de manera temprana, y realizar simulaciones de eventos de catástrofe para probar el proceso de recuperación de datos antes de que la aplicación se encuentre en productivo.
- Risk-008: Suscribirse a grupos de colaboración relacionados a la nueva tecnología para estar informado constantemente de los cambios que pueden llegar a suceder en el aspecto normativo con el objetivo de ganar tiempo para poder adecuarse a las nuevas reglamentaciones.
- Risk-005: Realizar capacitación en seguridad de la herramienta EC2 brindada por AWS para asegurar una correcta configuración del servidor y evitar problemas antes del despliegue en el ambiente de producción.
- Risk-004: Analizar balanceadores de cargas disponibles en AWS y realizar capacitación sobre buenas practicas de su configuración en el servidor de aplicaciones.

Conclusiones

Tomando como base mi experiencia profesional en el rubro de desarrollo y consultoría de software aplicada a la industria insurtech, trabajando como desarrollador y consultor en implementaciones tanto en Argentina como en Reino Unido, pude observar que los sistemas de gestión de documentos son una pieza clave para todas las organizaciones y se debe prestar especial atención en el diseño de los mismos, ya que de no asegurarse correctamente la información almacenada se pueden generar accesos no deseados o incluso pérdida de información, que en muchos casos puede tener datos sensibles de usuarios que deberían estar protegidos de manera adecuada. También pude notar, que cuando los volúmenes de información almacenada crecen de manera exponencial, los costos de infraestructura y mantenimiento aumentan considerablemente generando costos no previstos a las organizaciones.

Las problemáticas expuestas en el párrafo anterior, me hicieron pensar en la posibilidad de aplicar la tecnología blockchain a este tipo de soluciones con el objetivo de administrar de manera segura documentación digital aprovechando las ventajas que dicha tecnología brinda y para agregar valor al proyecto, exponer mediante reportes indicadores claves de información por medio de analítica de datos para dar soporte a la toma de decisiones.

La concreción del proyecto, me permitió cumplir con los objetivos planteados y desde el punto de vista personal, me llevó a aplicar muchos de los conocimientos adquiridos a lo largo de la carrera, así como también, investigar sobre tecnologías de las cuales tenía poco conocimiento.

Para concluir entiendo que este proceso fue una experiencia enriquecedora, con la cual pude mejorar considerablemente mis habilidades técnicas lo cual me permite estar mejor preparado para futuros desafíos.

Demo

En el siguiente link se puede encontrar un instructivo que explica como instalar y desplegar el prototipo de manera local.

Dentro del mismo directorio también se puede encontrar un archivo comprimido con el código fuente del mismo.

<https://drive.google.com/drive/folders/1wSSCH75m1FD4HHpHGUnXzouMdk9ii6pM?usp=sharing>

Referencias

Capterra. (17 de mayo de 2018). *Benefit of Blockchain, Part 1 Data Storage*. Recuperado de <https://blog.capterra.com/it-benefits-of-blockchain-data-storage/>

Document Logistix. (1 de mayo de 2021). *The History of Document Management System*. Recuperado de <https://www.document-logistix.com/the-history-of-document-management-from-filing-cabinet-to-cloud.html>

Filecoin Miner Analytics. (27 de abril de 2021). *Filecoin Average storage cost report*. Recuperado de <https://file.app/>

Gartner. (27 de abril de 2021). *Magic Quadrant or content Services Platforms*. Recuperado de <https://www.gartner.com/doc/reprints?id=1-1XPZSHJF&ct=191104&st=sb>

Gomez, M. (2021). *Test Driven Development*. Recuperado de <https://softwarecrafters.io/javascript/tdd-test-driven-development>

Poder Legislativo Nacional. (2001). *Ley 25.506 - Administración Financiera y de los Sistemas de Control Interno*. Argentina.

Poder Legislativo Nacional. (2018). *Ley 27.446 – Simplificación y Desburocratización de la administración pública*. Argentina.

Sitio oficial Alfresco. (2021). *Capacidades*. Recuperado de <https://www.alfresco.com/es/capacidades>

Sitio oficial Amazon. (2021). *Amazon S3*. Recuperado de <https://aws.amazon.com/es/s3>

Sitio oficial Atlassian. (2021). *Gitflow flujo de trabajo*. Recuperado de <https://www.atlassian.com/es/git/tutorials/comparing-workflows/gitflow-workflow>

Sitio oficial Box. (2021). *Overview*. Recuperado de <https://www.box.com/es-419/overview>

Sitio oficial Git. (2021). *About*. Recuperado de <https://git-scm.com/about>

Sitio oficial Filecoin. (2021). *How filecoin Works*. Recuperado de <https://docs.filecoin.io/about-filecoin/how-filecoin-works>

Sitio oficial MongoDB. (2021). *Advantages of MongoDB*. Recuperado de <https://www.mongodb.com/advantages-of-mongodb>

Sitio oficial Opentext. (2021). Recuperado de <https://www.opentext.com/products-and-solutions/products/enterprise-content-management>

Sitio oficial Pixelplex. (2021). *DocFlow: Intuitive Blockchain-Powered Document Management System*. Recuperado de <https://pixelplex.io/doc-flow/>

Sitio oficial de Scrum. (2021). *Helping People and Teams Solve Complex Problems*. Recuperado de <https://www.scrum.org/about>

Anexos

Anexo I - Leyes

Ley 25506 de Firma digital

Ley 27446 de Simplificación y Desburocratización de la administración pública.

Anexo II - Significado de acrónimos y códigos utilizados

PC: Personal Computer (Computadora Personal).

LAN: Local Area Network (Red de área local).

EDMS: Electronic Document Management System (Sistema de gestión de documentación digital).

SAAS: Software as a service (Software como servicio)

DMS: Document Management System (Sistema de gestión de documentos).

TI: Tecnología de la información.

GB: Giga bytes.

HTML: HyperText Marckup Language (Lenguaje de marcado para elaborar páginas web).

CSS: Cascading Style Sheets (Hoja de estilos en cascada).

JSON: Javascript Object Notation Object (Notación de objeto en javascript).

JSON-RPC: Stateless, light-weight remote procedure call (Protocolo de llamada a procedimiento remoto codificado en formato JSON).

BSON: Binary JSON (Notación de archivo binario en javascript).

PoST: Proof of space-time (Prueba de espacio tiempo)

ECM: Enterprise Content Management. (Gestor de contenido empresarial).

API: Aplication Programming Interface (Interfaz de programación de aplicaciones).

TDD: Test Driven Development (Desarrollo guiado por pruebas).

SSN: Súper Intendencia de Seguros de la Nación.

REST: Representational State Transfer (Transferencia de estado representacional).

TFG: Trabajo Final de Graduación.

DB: Database (Base de datos).

JWT: JSON Web Token.

IDP: Identity Provider. (Proveedor de identificaciones).

OIDC: Open Standard and Decentralized Authentication Protocol (Estándar de identificación digital descentralizado).

RTO: Recovery Time Objective (Tiempo objetivo de recuperación).

RPO: Recovery Point Objective (Punto objetivo de recuperación).

Anexo III - Comparación de costos de servicios de almacenamiento.

Ilustración 21 - Comparativa de Costos Filecoin vs S3

Costo promedio de almacenamiento			
Los cálculos se basan en un FIL-epoch que es de 30 segundos.			
Cálculo			
Costo de Filecoin por byte * 1073741824 = 2878048.971111595 attoFIL por GiB			
1 GIB POR DÍA	1 GIB POR MES 30 DIAS	1 GIB POR AÑO	COSTE MEDIO DE LA TRANSACCIÓN
0.000000008288781036801394 FIL ≈ \$ 0.0000005 USD	0.00000024866343110404184 FIL ≈ \$ 0.0000142 USD	0.000003025405078432509 FIL ≈ \$ 0.0001731 USD	0.00000000027707523375648562 FIL ≈ \$ 1.585701562788367e-9 USD
Comparación con Amazon S3: acceso poco frecuente			
Determinar el porcentaje de cuánto más barato o caro se compara el almacenamiento de Filecoin con Amazon S3 - Nivel de acceso poco frecuente. Ese nivel cuesta \$ 0.0134217728 por GiB por mes . Amazon recomienda este nivel de precios para datos de larga duración pero a los que se accede con poca frecuencia que necesitan acceso en milisegundos.			
Cálculo			
Costo de Filecoin / costo de Amazon = 0.11% del costo de Amazon S3			
1 GIB POR DÍA	1 GIB POR MES (30 DÍAS)	1 GIB POR AÑO	
0.11% del costo de Amazon S3	0.11% del costo de Amazon S3	0.11% del costo de Amazon S3	

Fuente: Extraído de file.app