

ESCUELA SUPERIOR POLITECNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

Desarrollo de interfaz gráfica para la limpieza y carga de archivos de datos de
producción de proteína animal

PROYECTO INTEGRADOR

Previo la obtención del Título de:

Ingeniero en Ciencias Computacionales

Presentado por:

Alexander Alzate Quintero

GUAYAQUIL – ECUADOR

Año: 2023

DEDICATORIA

El presente proyecto lo dedico a:

Mis padres, que siempre estuvieron dándome su apoyo incondicional y consejos que me ayudaron a llegar hasta este momento, convirtiéndome en una mejor persona y profesional

Mis hermanos por sus palabras de aliento y compañía.

Mi novia por acompañarme y escucharme siempre con una sonrisa en su rostro, dándome palabras de aliento y consejos que me ayudaron a superar grandes problemas que se presentaron en el camino.

AGRADECIMIENTOS

Agradezco a Dios por haberme bendecido con tan maravillosa familia, que siempre han creído en mí y me han acompañado en todo lo que he decidido hacer. Agradezco a mis padres que con su esfuerzo, dedicación y sacrificio me brindaron educación y valores.

Gracias a cada maestro que formó parte de mi desarrollo profesional, que me permitió dejar como recuerdo y legado este proyecto, que aportará al desarrollo y conocimiento de las generaciones futuras.

Gracias a ti lector, por permitir que mis investigaciones, conocimiento y esfuerzo formen parte de tu información mental.

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; Alexander Alzate Quintero doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”



Alexander Alzate Quintero

EVALUADORES

Ph.D. Lucia Marisol Villacres

Falconi

PROFESOR DE LA MATERIA

Msc. Allan Roberto Avendaño

Sudario

PROFESOR TUTOR

RESUMEN

Los retos que presentan las empresas pecuarias para tecnificar sus procesos y competir en un mercado en crecimiento e innovación constante las ha llevado a recurrir al análisis de datos. Ya que tienen procesos que generan grandes cantidades de datos. Datos de los cuales buscan sacar provecho, con el objetivo de identificar tendencias y patrones para aportar significativamente en la toma de decisiones. Sin embargo, para incorporar servicios de analítica de datos en sus procesos productivos se enfrentan a retos relacionados con la consistencia en los formatos de almacenamiento, el procesamiento de datos históricos y el manejo de estándares institucionales. Además, la inexistencia de herramientas que brinden soluciones simples a estos retos ha causado que solucionarlo sea costoso, dado que requiere de intervención profesional. Por lo que se desarrolló una interfaz gráfica para la limpieza y carga de archivos de datos de producción de proteína animal que puede ser usada por usuarios no técnicos. El alcance del diseño del proyecto fue definido por medio del análisis de requerimientos de usuario. Donde se decidió por una arquitectura sin servidor utilizando servicios de Amazon Web Services para su implementación. Como resultado de las evaluaciones del Plan de Observación y el formulario de Retroalimentación se conoce que la aplicación es considerada rápida y fácil de usar para los usuarios. Se concluye que la aplicación reduce el tiempo y complejidad del procesamiento de archivos de datos, a un bajo costo y fácil implementación de nuevas funcionalidades a las empresas pecuarias y de analítica.

Palabras Clave: Analítica, datos, empresas pecuarias, no técnicos, interfaz.

ABSTRACT

The challenges that livestock companies face in order to technify their processes and compete in a constantly growing and innovating market have led them to resort to data analysis. Since they have processes that generate large amounts of data. Data from which they seek to take advantage, with the aim of identifying trends and patterns to contribute significantly in decision making. However, to incorporate data analytics services in their production processes they face challenges related to consistency in storage formats, historical data processing and management of institutional standards. In addition, the lack of tools that provide simple solutions to these challenges has caused the solution to be costly, since it requires professional intervention. The present project seeks to develop a graphical interface for cleaning and loading animal protein production data files that can be used by non-technical users. The scope of the project design was defined through user requirements analysis. Where it was decided for a serverless architecture using Amazon Web Services for its implementation. As a result of the evaluations of the Observation Plan and the Feedback form it is known that the application is considered fast and easy to use for the users. It is concluded that the application reduces the time and complexity of data file processing, at a low cost and easy implementation of new functionalities.

Keywords: Data analysis, intuitive, livestock companies, non-technical, serverless.

Tabla de contenido

RESUMEN	I
ABSTRACT	II
1. INTRODUCCIÓN	1
1.1. DESCRIPCIÓN DEL PROBLEMA	1
1.2. JUSTIFICACIÓN DEL PROBLEMA	2
1.3. OBJETIVOS	3
1.3.1. <i>Objetivo general</i>	3
1.3.2. <i>Objetivos específicos</i>	3
1.4. MARCO TEÓRICO.....	3
1.4.1. <i>Analítica de datos</i>	3
1.4.2. <i>Big data</i>	4
1.4.3. <i>Empresas de producción de proteína animal</i>	5
1.4.4. <i>Experiencia de usuario</i>	5
1.4.5. <i>Asimetrix</i>	6
2. METODOLOGÍA	7
2.1. FASE 1: ANÁLISIS	7
2.1.1. <i>Usuarios de la solución</i>	9
2.2. FASE 2: REQUERIMIENTOS	9
2.2.1. <i>Funcionales</i>	9
2.2.2. <i>No funcionales</i>	11
2.2.3. <i>Alcance de la solución</i>	12
2.2.4. <i>Riesgos y beneficios de la solución</i>	13
2.3. DISEÑO.....	15
2.3.1. <i>Software</i>	15
2.4. PROTOTIPO.....	21
2.5. EVALUACIÓN	24
3. ANÁLISIS Y RESULTADOS	29
3.1. VALIDACIÓN	29
3.1.1. <i>Etapa de Observación</i>	29
3.1.2. <i>Etapa de Retroalimentación</i>	30
3.2. CASOS DE PRUEBA	31
3.2.1. <i>Usuarios</i>	31
3.2.2. <i>Proceso de evaluación</i>	32

3.3.	RESULTADOS	33
3.3.1.	<i>Resultados etapa de observación</i>	33
3.3.2.	<i>Resultados etapa de Retroalimentación</i>	38
3.4.	COSTOS	42
4.	CONCLUSIONES Y RECOMENDACIONES	45
4.1.	CONCLUSIONES	45
4.2.	RECOMENDACIONES.....	46
	BIBLIOGRAFÍA	47
	ANEXOS	48

ÍNDICE DE ILUSTRACIONES

ILUSTRACIÓN 1 DISEÑO DE ARQUITECTURA DE SOFTWARE – ARQUITECTURA SIN SERVIDOR	16
ILUSTRACIÓN 2 DIAGRAMA DE FLUJO DE PROCESO.....	17
ILUSTRACIÓN 3 DIAGRAMA DE CASOS DE USO.....	19
ILUSTRACIÓN 4 DIAGRAMA DE GANTT.....	20
ILUSTRACIÓN 5 PROTOTIPO CARGA DE DATOS.....	21
ILUSTRACIÓN 6 PROTOTIPO VARIABLES.....	22
ILUSTRACIÓN 7 PROTOTIPO HISTORIAL	23
ILUSTRACIÓN 8 PROTOTIPO BAJO NIVEL - SECCIÓN HOME	25
ILUSTRACIÓN 9 PROTOTIPO BAJO NIVEL - SECCIÓN DATOS.....	26
ILUSTRACIÓN 10 PROTOTIPO BAJO NIVEL - SECCIÓN DATOS (PROCESADOS).....	27
ILUSTRACIÓN 11 PLAN DE OBSERVACIÓN (AUTORÍA PROPIA)	30
ILUSTRACIÓN 13 TIEMPO EMPLEADO POR USUARIO PARA CARGAR UN ARCHIVO	34
ILUSTRACIÓN 14 TIEMPO EMPLEADO POR USUARIO PARA DESCRIBIR UN ARCHIVO	35
ILUSTRACIÓN 15 TIEMPO QUE TARDA EL SISTEMA EN PROCESAR UN ARCHIVO POR USUARIO	36
ILUSTRACIÓN 16 TIEMPO QUE TARDAN LOS USUARIOS EN ENCONTRAR EL ARCHIVO QUE CARGARON	37
ILUSTRACIÓN 17 TIEMPO QUE TARDAN LOS USUARIOS EN CONFIGURAR LA APLICACIÓN PARA QUE SOLICITE NUEVOS VALORES	38
ILUSTRACIÓN 18 ROLES DE USUARIOS EVALUADOS	39
ILUSTRACIÓN 19 USUARIOS QUE INTERACTUAN CON LA APLICACIÓN POR PRIMERA VEZ	39
ILUSTRACIÓN 20 FACILIDAD DE USO DE LA APLICACIÓN.....	40
ILUSTRACIÓN 21 EXPECTATIVAS SOBRE ASPECTO PROFESIONAL, VISUAL Y UTILIDAD	41
ILUSTRACIÓN 22 FORMULARIO GOOGLE PARTE 1.....	1
ILUSTRACIÓN 23 FORMULARIO GOOGLE PARTE 2.....	2
ILUSTRACIÓN 24 FORMULARIO GOOGLE PARTE 3.....	2
ILUSTRACIÓN 25 DATOS CAPTURADOS PLAN DE OBSERVACIÓN.....	3

ABREVIATURAS

ESPOL	Escuela Superior Politécnica del Litoral
AWS	Amazon Web Services

Capítulo 1

1. Introducción

1.1. Descripción del problema

Con el constante avance tecnológico las empresas se ven en la necesidad de tecnificar sus procesos para tener relevancia y competir en un mercado en crecimiento e innovación tecnológica continua. En dicho mercado los procesos productivos generan grandes cantidades de datos, el sacar provecho de estos datos es crucial para la mejora y crecimiento de la compañía.

De la necesidad de sacar ventaja de los datos generados en los procesos productivos surge la analítica de datos, que busca identificar tendencias y patrones para aportar significativamente en la toma de decisiones. Las empresas de analítica proveen servicios que mediante el proceso de exploración, transformación y examinación de los datos generan descubrimientos que ayudan a optimizar el rendimiento de procesos, obtener una mayor rentabilidad y mejorar las estrategias (datos M. p., 2019). En particular, las empresas de producción se benefician de la analítica de datos en la toma de decisiones para el crecimiento económico y la optimización de los procesos de producción.

Por otro lado, la incorporación de la analítica de datos en sus procesos productivos enfrenta retos relacionados con la consistencia en los formatos de almacenamiento de datos, el procesamiento de datos históricos y el manejo de estándares institucionales (El negocio de los datos, 2020). Las complicaciones para iniciar servicios de analítica de datos son una barrera para que nuevas empresas decidan embarcarse en este tipo de procesos.

1.2. Justificación del problema

Las empresas productoras de proteína animal o empresas pecuarias tecnifican sus procesos productivos con el objetivo de optimizar tareas y procedimientos para obtener mejores resultados, lo que implica realizar cambios estructurales para acoplarse a los requerimientos que estos exigen. Los cambios estructurales incluyen la adquisición de computadoras, sensores, equipos de telemetría, acceso a internet en las zonas de producción y demás herramientas tecnológicas que permitan la captura de datos y mayor control sobre cada uno de los procesos productivos.

Las empresas pecuarias no cuentan con las herramientas necesarias para el correcto procesamiento de las grandes cantidades de datos que generan. Por lo que las empresas de analítica ofrecen sus servicios de análisis donde utilizan los datos generados por las empresas pecuarias de manera inteligente con el objetivo de identificar patrones y tendencias que permitan optimizar los procesos productivos. Por consiguiente, las empresas pecuarias deben otorgar total acceso a los datos productivos de su compañía para ser almacenados en los sistemas de base de datos de la empresa de analítica.

Las empresas de producción encuentran problemático tener que estandarizar sus formatos de captura de datos para poder adquirir servicios de analítica debido a que es un proceso costoso que además repercute de manera negativa en la estructura de análisis actual de la compañía. Formatos donde se capturan datos por necesidades internas del negocio y que tendrían que excluir para acoplarse a la arquitectura de los sistemas de la empresa de analítica.

Asimetrix S.A. provee servicios de analítica de datos a productores de proteína animal, que busca romper barreras con el objetivo de atraer nuevos clientes, reduciendo el impacto inicial que tienen al adquirir los servicios que se ofrecen, por lo cual se plantea el desarrollo de una interfaz gráfica que permita a los usuarios procesar archivos sin importar su formato,

seleccionando las columnas o filas estrictamente necesarias para el correcto funcionamiento de los servicios provistos por Asimetrix.

1.3. Objetivos

1.3.1. Objetivo general

Crear una interfaz gráfica para ETL que permita la carga, descripción y procesamiento de archivos de datos en hojas de cálculo sin importar su estructura y/o formato por medio de la web para ser compatibles con los sistemas de las empresas de analítica.

1.3.2. Objetivos específicos

Desarrollar una interfaz gráfica en la web que permita seleccionar filas o columnas de un archivo de datos en hojas de cálculo para la manipulación de los datos.

Implementar un mecanismo de almacenamiento para uso de la empresa de analítica que sea compatible con los datos de las empresas de producción.

Implementar un mecanismo en la interfaz web para la verificación del estado de los datos posterior a la limpieza y carga de archivos de datos.

1.4. Marco teórico

1.4.1. Analítica de datos

La analítica de datos utiliza los datos en formato digital producidos en cualquier organización con el objetivo de generar conocimiento mediante el uso de modelos matemáticos, herramientas metodológicas y modelos estadísticos facilitando así la toma de decisiones de los empresarios. Los datos son valores cualitativos o cuantitativos (continuos o discretos) que, además de ser una inestimable herramienta de trabajo, se han convertido en un meritorio recurso capaz de proporcionarnos información (datos A. d., 2019). Para comprender la analítica de datos es necesario conocer que existen dos tipos de análisis de datos, el análisis predictivo o supervisado y el análisis no supervisado (datos A. d., 2019). El análisis predictivo o supervisado genera indicadores que se usan para comparar con la realidad a partir de variables objetivas; mientras que el análisis no supervisado no depende de una sola variable ni se puede determinar qué es lo que se analiza por lo que suele utilizarse para segmentar poblaciones. Para las empresas es prioridad analizar los procesos comerciales, pero la analítica no está restringida solo a esa a esa actividad (datos A. d., 2019).

1.4.2. Big data

Debido al crecimiento en los volúmenes de datos que se generan diariamente en los sistemas y actividades de la sociedad ha surgido la necesidad de administrar, generar métodos de almacenamiento y procesamiento de datos para suplir las fallas que tienen las bases de datos tradicionales (world, 2016). Por este motivo surge Big Data que su función principal es administrar grandes volúmenes de datos que se generan rápidamente y/o provienen de múltiples fuentes con mayor velocidad de acceso y procesamiento de los datos. Con el surgimiento del Big Data nace la ciencia de datos que engloba un conjunto de técnicas para el tratamiento y gestión de datos masivos con una perspectiva estadística. Estos beneficios hacen de Big Data una tecnología muy útil para el almacenamiento de datos masivos como los generados por las empresas de producción de proteína animal (Big Data: una exploración de investigaciones, 2017).

1.4.3. Empresas de producción de proteína animal

La función principal de las empresas de producción es transformar factores productivos con el fin de generar bienes y/o servicios que estén disponibles para el consumo a gran escala. Es considerada la unidad económica básica de la producción, siendo la encargada de producir y hacer llegar al consumidor bienes y servicios. Las empresas de producción de proteína animal o empresas pecuarias son compañías que ejercen en las áreas pecuarias y agrícolas. Acorde al instituto interamericano de cooperación para la agricultura se estima que la población mundial llegará a los 10 mil millones de personas para el 2050, lo que generará retos sociales como pobreza, inseguridad alimentaria y escasez de recursos (Occidental, 2021). Por lo que el sector pecuario tiene gran relevancia en su rol como generador de producción sustentable de alimentos (competitividad, 2020). Por ello es necesario que las empresas busquen tecnificar sus procesos de producción con el objetivo de optimizar sus procesos productivos y generando así una mayor producción (producción).

1.4.4. Experiencia de usuario

Vivir en una sociedad donde cada vez hay que entender mejor la tecnología y que se ha ido incluyendo en cada una de las actividades de nuestro diario vivir ha generado la necesidad de mejorar y facilitar la interacción entre el usuario y las interfaces tecnológicas; interacción que requiere de profesionales que se dedican a la experiencia de usuario cuyo rol es hacer que la tecnología sea amigable, fácil de usar, satisfactoria y útil (organizaciones, 2020). Entre los atributos de la experiencia de usuario destacan la usabilidad y la accesibilidad que son indispensables para el desarrollo de interfaces de usuario (metodos, 2015). La usabilidad es uno de los atributos de calidad de un producto, que analiza la facilidad de uso de una interfaz o tecnología, de igual forma la accesibilidad mide que tan posible es que una tecnología sea usada por la mayor cantidad de personas sin importar las limitaciones de cada individuo. Existen más conceptos que profundizan en el diseño centrado en usuario, pero todos con el fin de mejorar la relación entre usuarios y tecnologías (metodos, 2015).

1.4.5. Asimetrix

Asimetrix es una empresa global de datos agrícolas que ofrece soluciones y servicios de analítica de datos a productores de proteína animal. Usando computación en la nube, sensores y aprendizaje automático para mejorar la productividad y la efectividad en las granjas.

Capítulo 2

2. Metodología

Este capítulo presenta una descripción detallada del proceso utilizado para el desarrollo de la solución que está compuesto por 3 fases: Análisis, Requerimientos y Diseño.

En la fase de análisis se detallan las actividades realizadas para el levantamiento de requerimientos. Actividades como entrevistas o reuniones con el cliente y usuarios que utilizarán la solución.

En la fase de Requerimientos se detallan los requerimientos funcionales y no funcionales, alcance de la solución, beneficios y riesgos obtenidos como resultado de la fase 1.

En la fase de Diseño se describe la solución detalladamente, con gráficos como: Arquitectura del software, flujo de procesos, diagrama de identidades, planeación del tiempo de implementación, prototipos y evaluación.

2.1. Fase 1: Análisis

Se realizaron 4 reuniones con el cliente con una media de duración de 20 minutos cada una y 5 entrevistas con futuros usuarios de la solución con una duración de 15 minutos por persona. Las reuniones con el cliente y los usuarios permitieron recopilar información sobre las personas que utilizarán la solución brindando conocimiento sobre sus expectativas, experiencias con los métodos actuales y requerimientos funcionales.

Las personas que utilizarán la solución fueron determinadas en las reuniones con el cliente específicamente. Indicando que será utilizada por 3 tipos de personas: Una que sea la encargada de administrar la aplicación y realizar las configuraciones necesarias para su

correcto funcionamiento acorde al modelo de negocio, otra que se encargará de cargar archivos de datos, describirlos, enviarlos a procesar, administrarlos y, por último, una persona que solo pueda visualizar un listado de los archivos de datos que han sido cargados.

Se espera que la solución sea utilizada por personas de diferentes partes del mundo por lo que el idioma es de suma importancia para la ****(buscar palabra) de los futuros usuarios. Debido a la internacionalización de la solución se vuelve necesaria la disponibilidad de esta para las diferentes zonas desde las que se desee acceder. También se espera que la solución sea capaz de procesar archivos con grandes cantidades de datos sin interferencias ni lentitud, por lo que la escalabilidad y la tolerancia a fallos juegan un rol importante en su desarrollo.

Durante las reuniones y entrevistas con los futuros usuarios de la solución, comentaron sus experiencias con los procesos actuales. Los 5 entrevistados coincidieron con que el tiempo de espera medio para la carga de datos de archivos que no han sido procesados antes (primera vez que se cargan) es de alrededor de 1 semana y media. De igual manera indican que cuando hay cambios en los archivos hacer las correcciones es complejo para ellos. Esto debido a que para realizar los cambios es necesario tener conocimientos de programación y esa no es su área de conocimiento.

El cliente durante las reuniones hizo énfasis en el desarrollo de un proceso para la descripción y limpieza del archivo de datos de acuerdo con las instrucciones generadas por el usuario. Lo que permite que usuarios sin conocimientos en programación puedan realizar el procesamiento y carga de archivos de datos con mayor facilidad.

Como resultado de las entrevistas realizadas se determinó que para el desarrollo de la solución es recomendable utilizar una arquitectura de software sin servidor. Arquitectura que permite cubrir necesidades como la escalabilidad y disponibilidad del sistema.

2.1.1. Usuarios de la solución

A partir de las reuniones con el cliente y los usuarios fueron identificados los siguientes roles para la aplicación:

- **Usuario:** Este tipo de usuario cuenta con permisos para editar, eliminar y cargar nuevos archivos de datos. Además, visualizar el historial de archivos cargados.
- **Visualizador:** Este tipo de usuario cuenta con permisos de visualización.
- **Administrador:** Usuario que puede crear, modificar y eliminar configuraciones de carga como: categorías de datos y campos requeridos.

2.2. Fase 2: Requerimientos

Luego de las reuniones con el cliente y futuros usuarios de la solución se obtuvieron los siguientes requerimientos funcionales y no funcionales:

2.2.1. Funcionales

- **Usuario carga archivos de datos**

El rol usuario puede cargar archivos de cálculo en la interfaz donde se validará el tipo de archivo e indicará la orientación de los datos (Fila/columna).

- **Usuario describe las ubicaciones en el archivo de datos**

El rol usuario puede describir las columnas o filas del archivo cargado con el objetivo de indicar que datos corresponden a cada valor solicitado en la interfaz.

- **Usuario accede al listado de archivos de datos cargados**

El rol usuario puede observar una lista histórica de los archivos de datos que han sido cargados, descargarlos y visualizarlos.

- **Usuario elimina archivos históricos de datos cargados**

El rol usuario puede eliminar los archivos de datos que se han cargado en su compañía.

- **Administrador configura valores estrictamente necesarios**

El rol administrador permite al usuario configurar los valores que son estrictamente necesarios para el correcto funcionamiento de sus sistemas acorde a su modelo de negocio.

- **Administrador accede al listado de archivos cargados por compañía**

El rol administrador permite al usuario seleccionar una compañía y visualizar todos los archivos de datos cargados dentro de cada categoría.

- **Administrador elimina y carga archivos de datos**

El rol administrador permite al usuario eliminar y cargar archivos de datos dentro de cualquier compañía.

- **Visualizador observa listado de archivos de datos cargados**

El rol de visualizador permite al usuario solamente observar el listado de archivos de datos que han sido cargados en su compañía.

- **Visualizador descarga archivos de datos procesados**

El rol de visualizador permite al usuario descargar archivos de datos que han sido cargados en su compañía.

2.2.2. No funcionales

- **Internacionalización**

El sistema debe permitir incorporar lenguajes para traducción de idioma del contenido, principalmente español e inglés como idiomas principales.

- **Tolerancia a fallos**

El sistema debe ser capaz de prever problemas de fallos o bloqueos durante el proceso de descripción y transformación del archivo.

- **Validación**

El sistema debe validar los datos que descritos por el usuario tomando en consideración tipos de datos, consistencia de los datos e integridad de los datos.

- **Mantenimiento**

El diseño del sistema debe permitir la incorporación de nuevas funcionalidades o la corrección de errores con facilidad.

- **Escalabilidad**

El sistema debe ser capaz de procesar archivos con grandes cantidades de datos con tiempos de respuestas cortos. Debe ser capaz de acoplarse a las necesidades empresariales de cada usuario ofreciendo la misma rapidez de procesamiento.

2.2.3. Alcance de la solución

Para el desarrollo de la solución acorde a los requerimientos obtenidos en las entrevistas y encuestas se opta por separar la solución en 4 secciones: Interfaz web, API Gateway, funciones como microservicios y almacenamiento.

La interfaz web será el servicio que los usuarios utilizarán para la carga, descripción y visualización de archivos de datos. En la interfaz es necesario contar con una sección para la carga de los archivos y otra para visualizar el histórico de archivos cargados por el usuario en cada empresa. De igual manera permitir al usuario administrar sus archivos de datos sea cargando nuevos, eliminando o modificándolos. Todo esto ligado a los permisos otorgados por cada rol de usuario.

El API Gateway será el encargado de administrar el control de tráfico, autorizaciones y acceso de las solicitudes generadas en la interfaz web. Este servicio facilita la publicación, monitoreo y mantenimiento de endpoints o “punto final” actuando como puerta de entrada para las solicitudes de la interfaz web. El API Gateway será de tipo REST el cual permite cargas de trabajo sin servidor.

Las funciones como microservicios son scripts de código que se ejecutarán cada vez que sean llamados por el API Gateway. Estas funciones serán las encargadas de realizar todos los procesos necesarios para la limpieza, transformación y almacenamiento de los datos acorde a las instrucciones dadas por el usuario en la interfaz web. Estas tienen la

particularidad de que son reutilizables, de rápida implementación y compatibles con los diferentes proveedores de servicios en la nube.

Para el almacenamiento de los datos se decide utilizar una base de datos relacional y otra no relacional. La base de datos relacional será la encargada de almacenar todos los datos estrictamente necesarios para el correcto funcionamiento y compatibilidad de los servicios acorde al modelo de negocio. Por otra parte, la base de datos no relacional permitirá que los usuarios almacenen datos propios no ligados a la estructura del sistema, datos que varían de usuario en usuario.

En conclusión, el alcance de la solución contempla el procesamiento y carga de datos provenientes de hojas de cálculo sin importar su contexto de procedencia, es decir, es una solución útil para cualquier empresa que quiera procesar y limpiar sus datos almacenados en hojas de cálculo sin ser una empresa de producción de proteína animal. Teniendo en cuenta atributos como la rapidez de desarrollo, escalabilidad y mantenibilidad, propiedades esenciales para el correcto funcionamiento de la solución.

2.2.4. Riesgos y beneficios de la solución

De acuerdo con el alcance de la solución es necesario analizar los riesgos y beneficios que esta trae con su implementación.

2.2.4.1. Riesgos

Entre los principales riesgos detectados en la propuesta de la solución tenemos los siguientes: Formatos de almacenamiento de datos complejos y archivos de datos demasiado pesados.

Es muy probable que usuarios intenten cargar archivos de datos con formatos demasiado complejos que se salgan del requerimiento de ordenamiento filas o columnas. Debido a que

el proceso de asignación de variables no es compatible con archivos que contengan múltiples tablas o combinaciones de datos entre filas y columnas, ya que la solución no contempla dichos niveles de complejidad.

Carga de archivos de datos demasiados pesados que generen lentitud para el cliente al momento de renderizarlos. Este problema puede presentarse generalmente cuando las propiedades físicas del dispositivo en el que se está realizando el proceso de carga no sean suficientes para soportar archivos con pesos demasiado altos.

2.2.4.2. Beneficios

Los beneficios que brinda la solución sobre los procesos utilizados actualmente denotan la superioridad y necesidad de su pronta implementación. Entre los principales beneficios tenemos: Rapidez para la carga de nuevos archivos de datos, reducción de costos productivos y compatibilidad de los datos con el sistema de almacenamiento.

La solución permite que el proceso de limpieza, transformación y almacenamiento de los datos sea mucho más veloz y eficaz que los métodos utilizados actualmente. Brindando la facilidad a los usuarios de ellos mismos realizar la descripción de los archivos de datos y enviarlos a procesar sin la necesidad de tener conocimientos en programación. De igual forma para reprocesar archivos que han cambiado con respecto al tiempo podrán modificar la descripción de cada formato a necesidad.

Se reducen costos de producción debido a que se elimina del proceso la necesidad de uso de un programador para describir, transformar y almacenar los archivos de datos enviados por los usuarios. De igual manera la propuesta de solución utiliza la arquitectura sin servidor que entre sus principales beneficios está cobro por uso de recursos, es decir, se paga por lo que se usa.

Los usuarios podrán almacenar los datos necesarios para el modelo de negocio y datos de interés propio que varían entre usuarios. Cubriendo así la necesidad de las empresas de analítica de datos y las compañías productoras de proteína animal sin interferir en los requerimientos del modelo de negocio.

2.3. Diseño

En esta sección se describirá a detalle la estructura del sistema, cada uno de sus componentes y el motivo de su implementación. De igual forma se presentan gráficos y prototipos que brindan una mayor visión de su composición y apariencia.

2.3.1. Software

Para el diseño del software se decidió utilizar tecnologías en la nube por la arquitectura sin servidor. Para la implementación de la solución se utilizarán específicamente los servicios de Amazon Web Services (conocido también por sus siglas AWS).

2.3.1.1. Arquitectura

Para el diseño de solución se planteó utilizar arquitectura sin servidor o serverless, como se puede observar en la Ilustración 1. Una solución desarrollada en su mayor parte con servicios de AWS por su facilidad de implementación y costos de implementación.

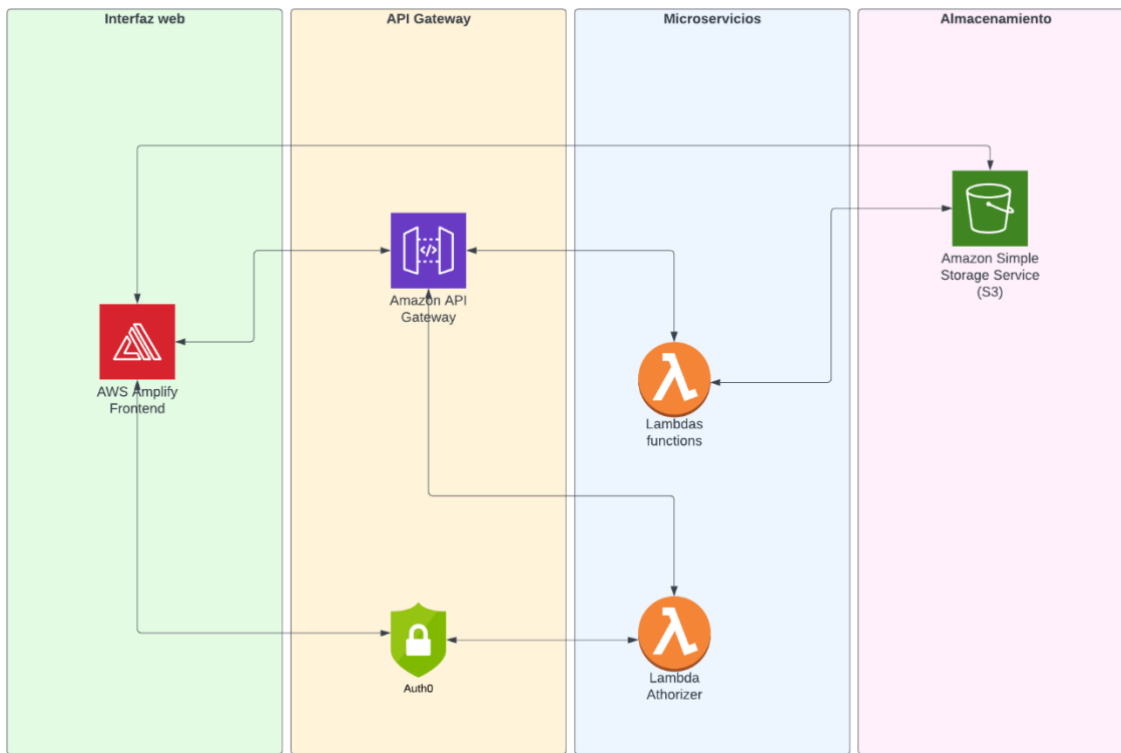


Ilustración 1 Diseño de arquitectura de software – Arquitectura sin servidor

AWS Amplify es un servicio de AWS que permite crear y alojar aplicaciones fácilmente, implementando otros servicios de la compañía que permiten evolucionar los proyectos a necesidad. No requiere un alto nivel de conocimiento sobre computación en la nube y consta dentro de la capa gratuita ofrecida por AWS.

AWS S3 es un servicio de AWS que permite almacenar objetos de manera segura y escalable. Este servicio será utilizado para almacenar los archivos de datos resultantes del proceso de limpieza y transformación. De igual manera almacenará los scripts de limpieza acorde a las descripciones dadas por el usuario.

AWS Lambda es un servicio de AWS que ejecuta código inicializado por eventos. Se utilizará para ejecutar todos los procesos backend necesarios para el procesamiento de los

archivos y funcionamiento general de la interfaz web. En particular existirá una Lambda Authorizer que se encargará de validar la autenticación y autorización de las solicitudes recibidas desde la interfaz web.

Auth0 es una plataforma que brinda servicios de autenticación y autorización con fácil implementación. Se utilizará este servicio para gestionar las credenciales de usuarios, roles y permisos dentro de la interfaz.

AWS API Gateway es un servicio de AWS que facilita la administración, monitoreo, desarrollo y seguridad de endpoints. Este servicio será el encargado de distribuir las solicitudes generadas en la interfaz web hacia cada una de las lambdas, además será el encargado del control de autenticación y autorizaciones.

2.3.1.2. Diagramas

Se utilizaron tres tipos de diagramas que permitieron identificar y visualizar mejor los procesos que conlleva la solución.

- **Diagrama del flujo del proceso de carga de archivos de datos**

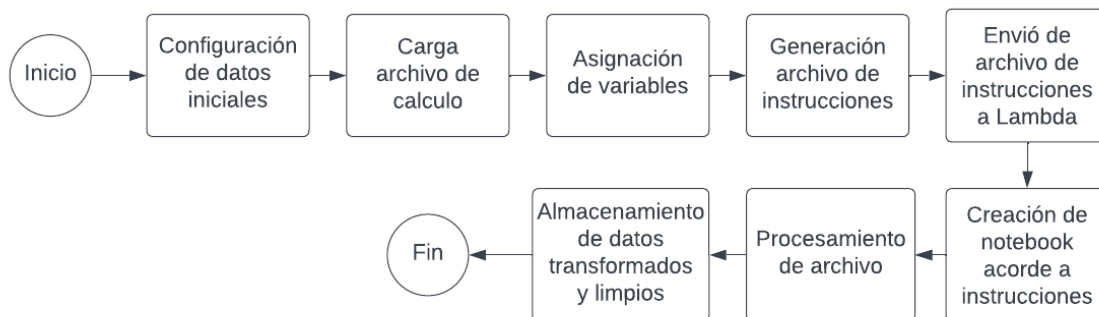


Ilustración 2 Diagrama de flujo de proceso

El diagrama de la Ilustración 2 brinda una visión pasó a paso del flujo de proceso que se realiza para la carga de archivos de datos. Proceso que inicia con la carga del archivo en la interfaz web y culmina con el almacenamiento de los datos ya transformados en S3.

- **Diagrama de casos de uso**

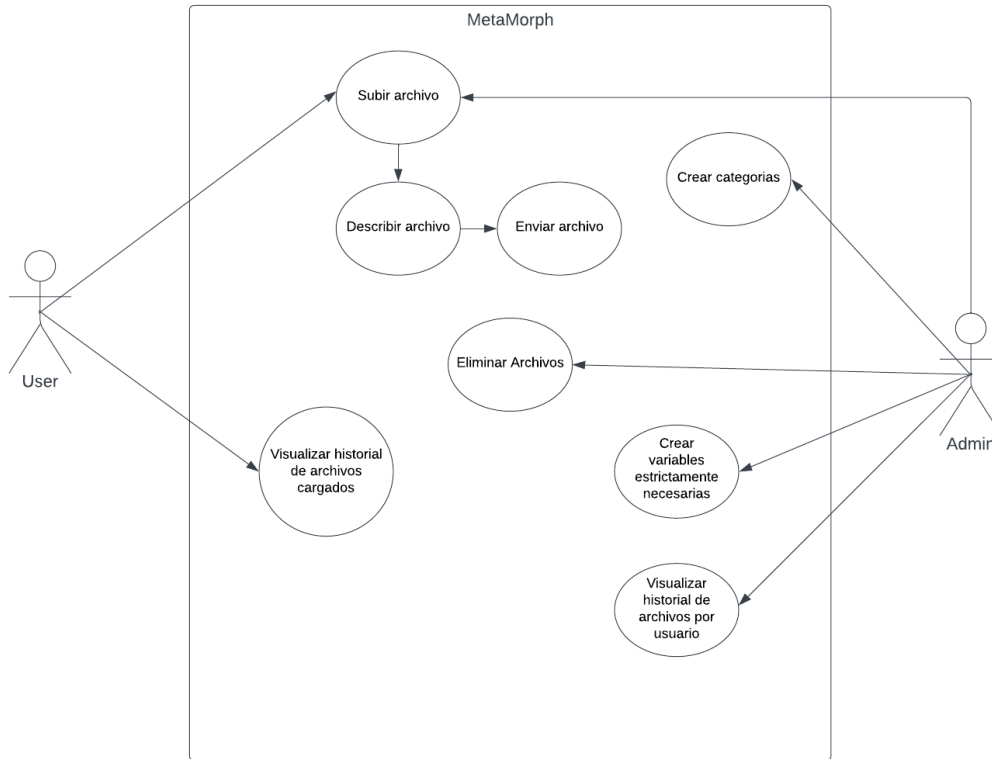


Ilustración 3 Diagrama de casos de uso

En el diagrama de la Ilustración 3 podemos visualizar todos los casos de uso que tiene cada rol de usuario. En este diagrama se puede visualizar cuáles son los procesos compartidos entre roles y sirvió para entender mejor las funcionalidades necesarias de cada rol. En el diagrama se visualizan los 3 roles definidos con anterioridad y las acciones que pueden realizar dentro de la interfaz.

- **Diagrama de Gantt**

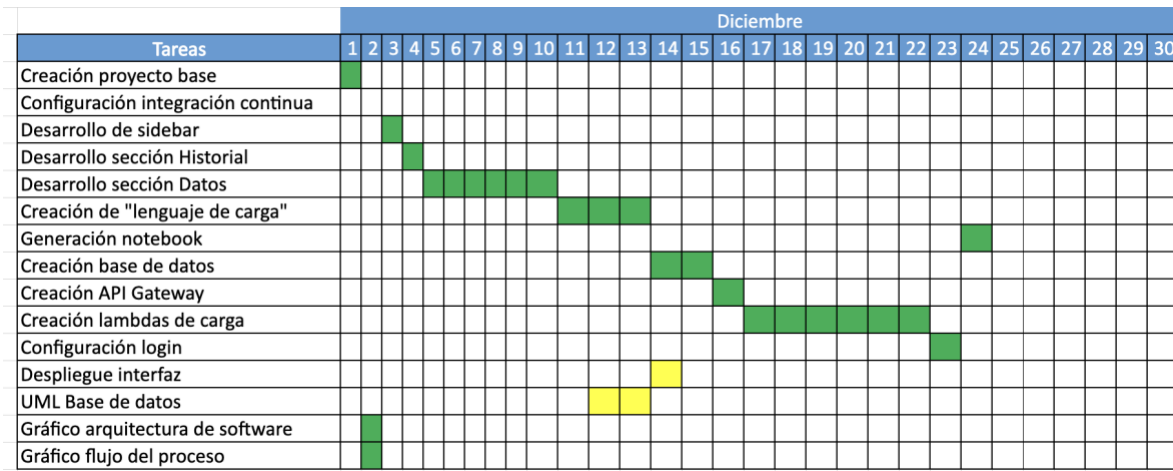


Ilustración 4 Diagrama de Gantt

El diagrama de la Ilustración 4 nos permitió planificar el tiempo de desarrollo de la solución en su totalidad en días tomando como referencia el mes de diciembre. Cada una de las tareas en el diagrama cuenta con el tiempo estimado de desarrollo en días. Los tiempos de color verde son los que hay que cumplir de forma obligatoria y los amarillos son sugeridos en esas fechas por la carga total del día.

2.4. Prototipo

Para el prototipado se tomó en cuenta lo solicitado por el cliente, estableciendo las pantallas de la interfaz que tienen mayor relevancia debido al objetivo principal de la aplicación.

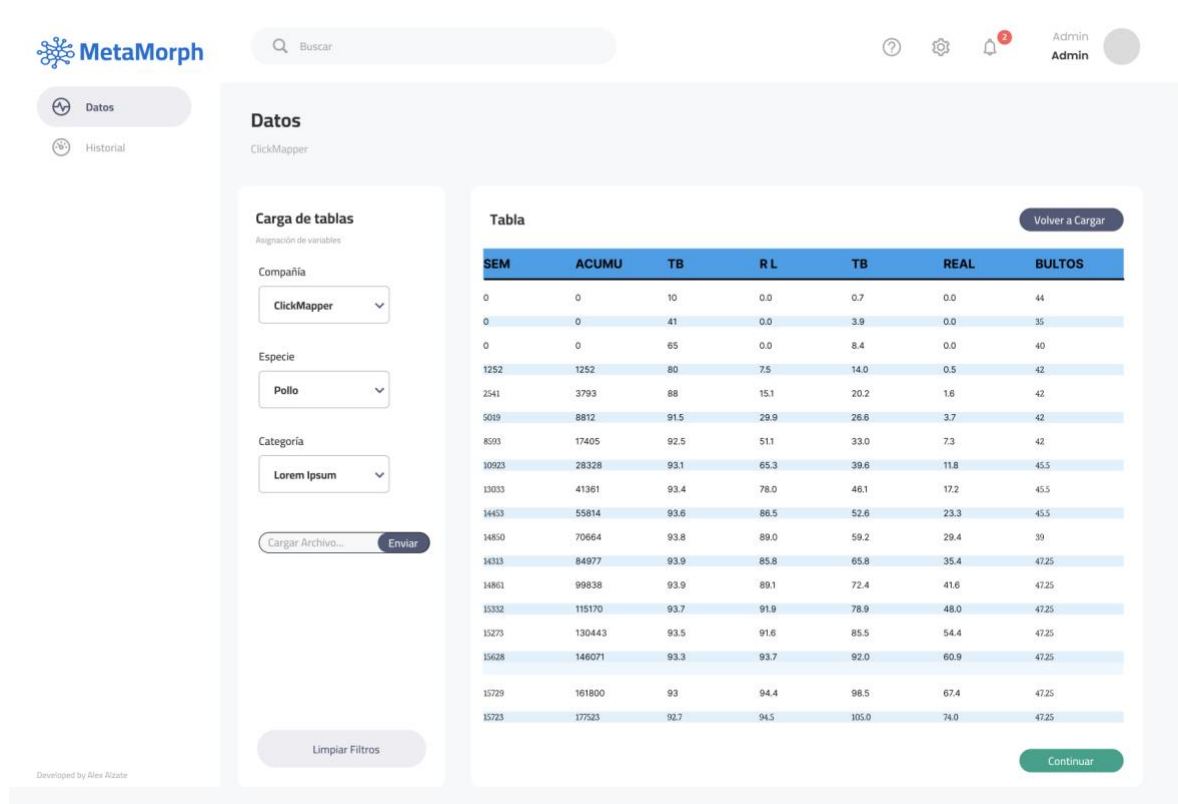


Ilustración 5 Prototipo carga de datos

En la Ilustración 5 se visualiza la sección de carga de datos que cuenta con un menú lateral y la sección de visualización del archivo de datos cargado. En el menú lateral se realiza la configuración de metadatos como la compañía, especie y categoría.

También se puede visualizar en la parte superior una barra de menú que cuenta con el perfil de usuario donde se visualiza el nombre, rol del usuario y foto.

The screenshot shows the MetaMorph application interface. At the top left is the MetaMorph logo. A search bar with the text 'Buscar' is located at the top center. On the top right, there are icons for help, settings, and a notification bell with a red '2' badge, along with the text 'Admin Admin'. Below the search bar, there are two tabs: 'Datos' (selected) and 'Historial'. The main content area is titled 'Datos' and 'ClickMapper'. On the left side, there is a 'Variables' section with the subtitle 'Asignación de variables'. It contains three categories: 'Geografía' (Geography) with variables like 'Granja', 'Lote', 'Cajón', 'País', and 'Departamento'; 'Alimentación' (Feeding) with variables like 'Consumo Diario', 'Consumo total', 'Engorde diario', 'Proveedor pollo', and 'Peso Pavo'; and 'Animal y Plantas' (Animal and Plants) with variables like 'Sexo', 'Especie', 'Estrato', 'Número de aves', and 'Etapas'. Below these is an 'Otros' (Others) section with variables like 'Fecha Entrada', 'Fecha Salida', 'Conversiones', 'Fase', and 'Línea Genética'. A 'Limpiar Filtros' button is at the bottom of the sidebar. On the right side, there is a 'Tabla' section with a 'Volver a Cargar' button. The table has columns: SEM, ACUMU, TB, R L, TB, REAL, and BULTOS. The table contains 20 rows of data. A 'Continuar' button is at the bottom right of the table.

SEM	ACUMU	TB	R L	TB	REAL	BULTOS
0	0	10	0.0	0.7	0.0	44
0	0	41	0.0	3.9	0.0	35
0	0	65	0.0	8.4	0.0	40
1252	1252	80	7.5	14.0	0.5	42
2541	3793	88	15.1	20.2	1.6	42
5019	8812	91.5	29.9	26.6	3.7	42
8593	17405	92.5	51.1	33.0	7.3	42
10923	28328	93.1	65.3	39.6	11.8	45.5
13033	41361	93.4	78.0	46.1	17.2	45.5
14453	55814	93.6	86.5	52.6	23.3	45.5
14850	70664	93.8	89.0	59.2	29.4	39
14313	84977	93.9	85.8	65.8	35.4	47.25
14861	99838	93.9	89.1	72.4	41.6	47.25
15332	115170	93.7	91.9	78.9	48.0	47.25
15273	130443	93.5	91.6	85.5	54.4	47.25
15628	146071	93.3	93.7	92.0	60.9	47.25
15729	161800	93	94.4	98.5	67.4	47.25
15723	17523	92.7	94.5	105.0	74.0	47.25

Ilustración 6 Prototipo variables

En la Ilustración 6 se puede observar una barra lateral y el archivo de datos cargado al lado derecho. La barra lateral muestra las variables que son estrictamente necesarias que permiten la compatibilidad con el sistema del modelo de negocio. En esta pantalla el usuario irá seleccionando cada una de las variables y la relacionará con la columna o fila del archivo de datos indicando el índice en el que se encuentra cada valor.

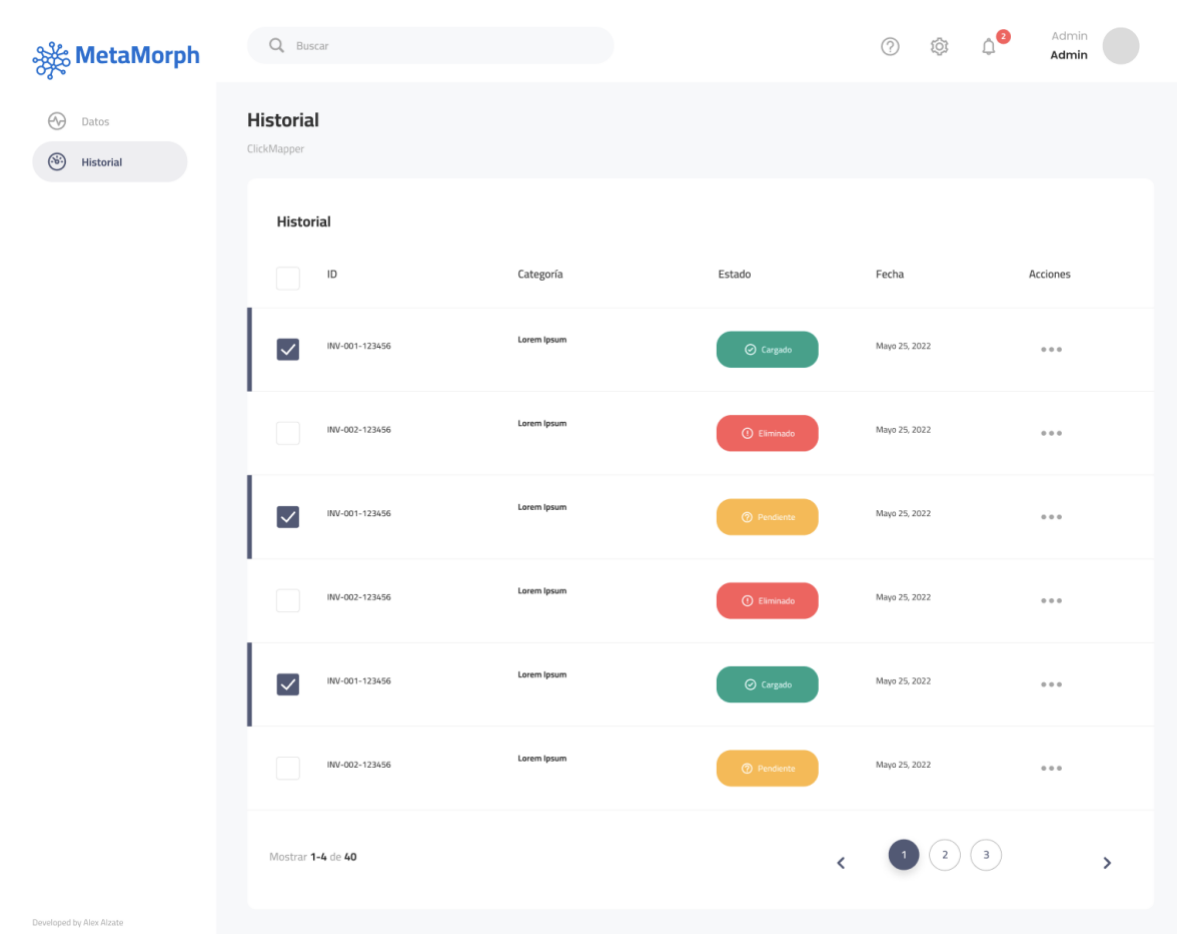


Ilustración 7 Prototipo historial

En la Ilustración 7 se visualiza la pantalla de historial. En esta pantalla se mostrará un listado de los archivos que han sido cargados por el usuario con rol usuario o administrador. Aquí se puede visualizar todo el histórico de datos lo que brinda la opción al usuario de descargar el archivo original y el procesado.

2.5. Evaluación

Para evaluar la evolución de los prototipos fue necesario realizar 3 reuniones con el cliente. Reuniones con una duración promedio de 15 minutos donde se obtuvo como resultado un primer prototipo de bajo nivel y un segundo prototipo de alto nivel que fue aceptado por el cliente en la última reunión.

- **Primer Prototipo**

El primer prototipo fue desarrollado en papel y fue esencial para comprender junto al cliente lo que esperaba de la solución y como debería funcionar.

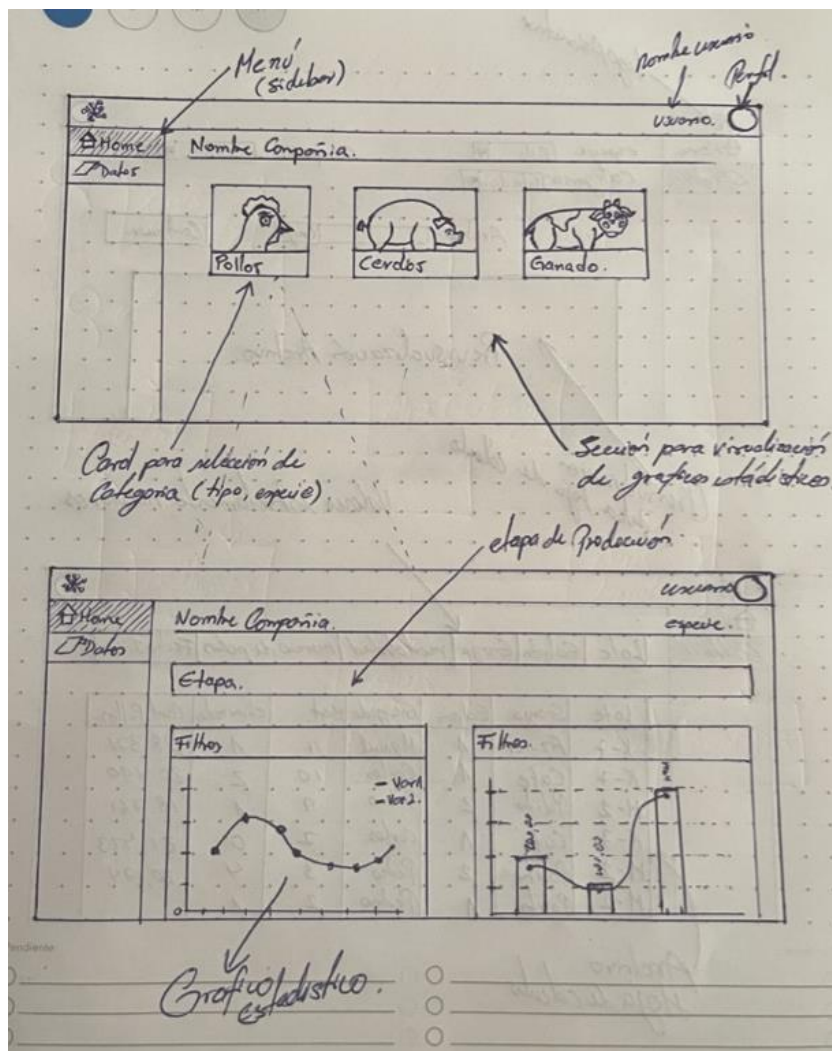


Ilustración 8 Prototipo bajo nivel - sección Home

En la Ilustración 8 se visualizan dos pantallas correspondientes a la sección Home o inicio. En estas pantallas se visualizaban los datos de los archivos cargados por medio de gráficos estadísticos. Estas pantallas contaban con filtros de búsqueda y categorías. Esta sección fue retirada del prototipo en la segunda reunión con el cliente, debido a que no era necesario que la solución presentara los datos de dicha forma ya que se cuenta con otras herramientas para ese propósito.

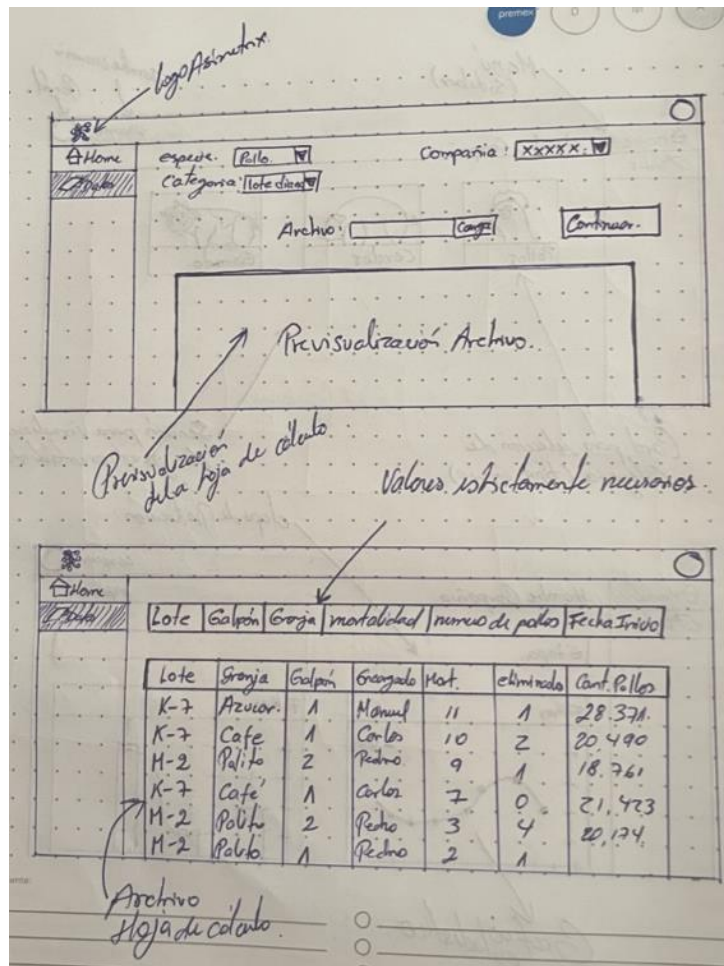
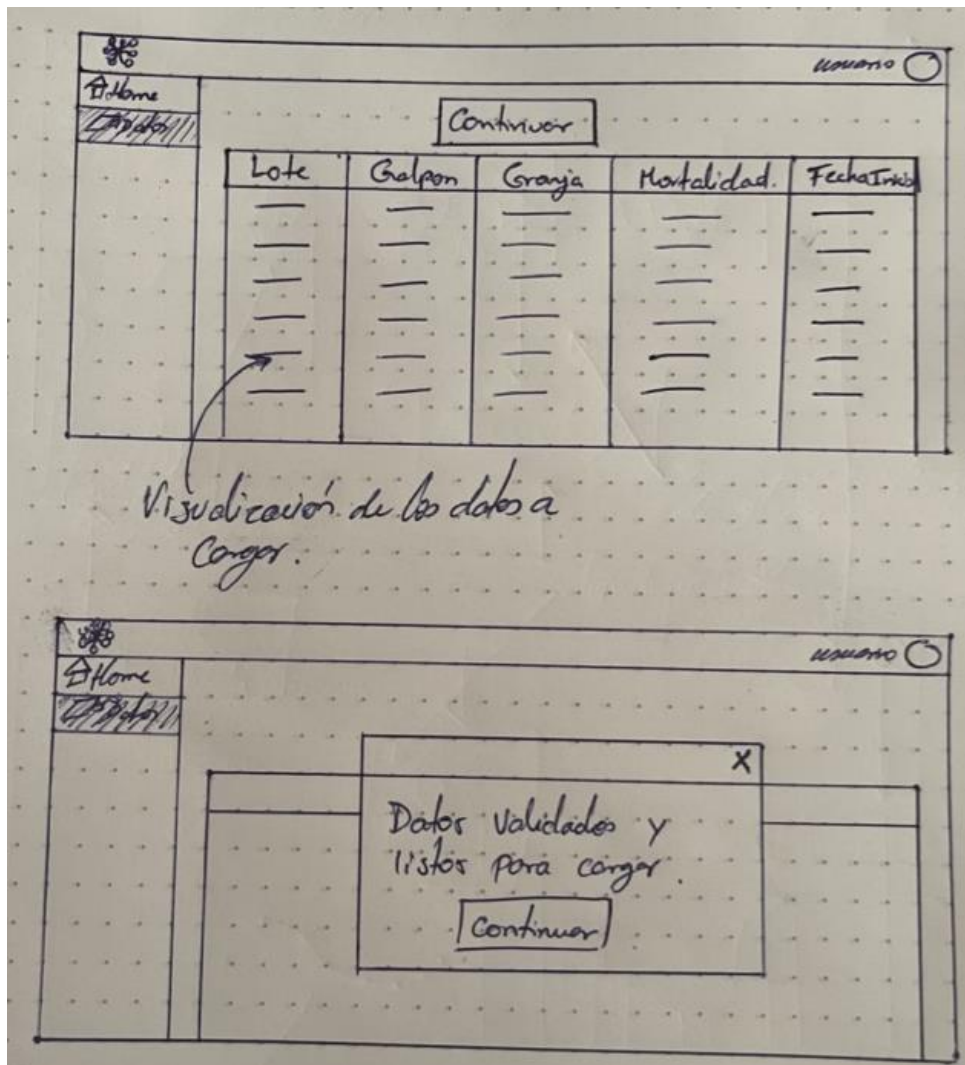


Ilustración 9 Prototipo bajo nivel - sección Datos

En la Ilustración 9 se visualizan dos pantallas de la sección Datos. En esta sección se visualiza el proceso de carga de los archivos de datos. En la parte superior de la primera pantalla se visualiza la configuración básica de meta data y en la parte inferior la previsualización del archivo de datos cargado.

En la pantalla de la parte de inferior de la Ilustración 9 se visualiza en la parte superior el listado de las variables estrictamente necesarias para el correcto funcionamiento del sistema acorde al modelo de negocio y en la parte inferior el archivo de datos renderizado para relacionar con las variables.



Visualización de los datos a cargar.

Ilustración 10 Prototipo bajo nivel - sección Datos (procesados)

En la Ilustración 10 se visualizan dos pantallas de la sección Datos consecuente de la Ilustración 9. En la pantalla de la parte superior se muestran los datos seleccionados y procesados acorde a las instrucciones dadas por el usuario previo a su almacenamiento. En la pantalla de la parte inferior se retroalimenta al usuario sobre el proceso de almacenamiento de los datos enviados.

- **Segundo Prototipo**

El segundo prototipo que se visualiza en la Ilustración 5 es la versión final y aceptada por el cliente como resultado de la tercera y última reunión. En esta versión se eliminaron los gráficos estadísticos y se añadió una sección nueva de historial como se visualiza en la Ilustración 7.

Este prototipo fue validado por 6 personas que forman parte del equipo de desarrollo de Asimetrix. Estos brindaron recomendaciones y su opinión con respecto a colores y distribución de los componentes en pantalla para la sección de carga. Consejos que fueron aplicados al prototipo y que permitieron cumplir con las expectativas del cliente.

Capítulo 3

3. Análisis y resultados

En esta sección se analizarán los procesos de validación, casos de prueba, resultados y costos de producción.

3.1. Validación

Para el proceso de validación con usuarios consta de dos etapas: etapa de observación y etapa de Retroalimentación. Para la etapa de observación se desarrolló un Plan de Observación con el fin de brindar una guía al evaluador del aplicativo sobre lo que debe observar en cada uno de los pasos realizados por el usuario evaluado. Mientras que la etapa de Retroalimentación se realiza al final de cada proceso permitiendo al usuario evaluado brindar su opinión respecto al uso del aplicativo.

3.1.1. Etapa de Observación

En la Etapa de Observación se utilizó como herramienta un Plan de Observación, que radica en desarrollar un formulario que sirve como guía para la captura de datos durante el proceso de evaluación de usuarios.

Plan de observación

Objetivos

Carga de Archivos

- Tiempo promedio
- Se realizó con éxito
- Nivel de complejidad
- Intentos

Búsqueda y visualización de archivos

- Tiempo promedio
- Se realizó con éxito
- Intentos
- Cantidad de errores generados

Descripción de Archivo

- Tiempo promedio
- Se realizó con éxito
- Intentos
- Cantidad de errores generados

Configuraciones variables

- Tiempo promedio
- Se realizó con éxito
- Intentos
- Cantidad de errores generados

Procesamiento

- Tiempo promedio
- Cantidad de errores generados

Proceso de observación

1. Sujeto de observación. – Personas entre 25 y 40 años
2. Ubicación. – Reunión virtual
3. Tiempo de espera promedio de evaluación. – 35 minutos
4. Patrones de comportamiento
 - a. Visualizaban el aplicativo minuciosamente tratando de comprender los datos que se le pedían.
 - b. Interactuaban con las configuraciones de visualización de la tabla cambiando las posiciones iniciales de los datos.
 - c. Descargaban el archivo procesado y el que cargaron para comparar y comprobar que se efectuaron cambios acordes a lo que describieron.
5. Objeto de observación
 - a. Proceso de carga y descripción de un documento de cálculo básico.

Ilustración 11 Plan de observación (Autoría Propia)

El Plan de Observación permite enfocar lo que se visualiza en cada uno de los procesos de evaluación con el usuario generando una toma de datos más organizada. Para ello se separó en dos secciones como se visualiza en la Ilustración 11, la primera sección consta de los objetivos a observar y los datos a capturar, y la segunda sección describe el proceso realizado durante la acción de observación.

3.1.2. Etapa de Retroalimentación

La Etapa de Retroalimentación se realiza al final de la evaluación mediante un formulario para capturar los comentarios del usuario sobre el proceso realizado. El formulario de

Retroalimentación se creó para analizar los aspectos negativos y positivos de los comentarios dados por los usuarios. Se decidió usar el formulario de Retroalimentación debido a que permite visualizar de manera ordenada las impresiones de los usuarios acorde al prototipo presentado.

El formulario de retroalimentación consta de 8 preguntas. Donde cada una de las preguntas abarca un área de opinión diferente mediante el cual guiamos los comentarios y opiniones del usuario evaluado con respecto a lo que necesitamos saber sobre la aplicación. El formulario completo puede visualizarse en el Anexo 1

3.2. Casos de Prueba

Para proceder con los casos de prueba se realizaron evaluaciones individuales con 18 usuarios, se utilizaron 3 herramientas durante las evaluaciones y el proceso de evaluación tuvo una duración de 30 minutos promedio.

3.2.1. Usuarios

Las pruebas se realizaron con 18 usuarios, 10 de manera individual y 8 de manera colectiva. El grupo evaluado individualmente se distribuye en 6 desarrolladores de software, 3 consultores y 1 gerente general.

Los usuarios desarrolladores de software serán los encargados de administrar y configurar la aplicación con el rol administrador en Asimetrix, por lo que evaluar el proceso de configuración con estos usuarios es de suma importancia para conocer sus necesidades y opiniones.

Los consultores son los usuarios más cercanos a las empresas de producción de proteína animal, son los encargados de acompañarlos durante todo el proceso de integración y uso

de los servicios de Asimetrix, por lo que es la categoría con mayor relevancia para la evaluación del aplicativo.

Entre los 18 usuarios evaluados había dos con conocimientos previos sobre el uso de la aplicación. Estos dos usuarios fueron los encargados de supervisar y validar cada una de las historias de usuario durante el proceso de desarrollo de la aplicación.

3.2.2. Proceso de evaluación

Para la ejecución del proceso de evaluación se agendaron reuniones virtuales de 30 minutos. Se explicó al usuario sobre el proceso de evaluación y se les envió 3 archivos para utilizar durante la prueba: Un archivo de calculo simple, un archivo de cálculo con una estructura compleja y un archivo JSON para realizar las configuraciones de la aplicación.

A cada usuario se aginó las credenciales del portal de Asimetrix, se creó una compañía de prueba y se dio acceso a la cuenta AWS de la compañía para las configuraciones de la aplicación.

Se indicó a los usuarios que debían cargar primero el documento de cálculo simple, realizando el proceso sin asistencia ni guía. Proceso que tuvo como objetivo observar el comportamiento del usuario al interactuar por primera vez con la interfaz e ir tomando apuntes acordes a lo indicado en el plan de observación.

Para cargar el documento complejo es necesario crear un nuevo tipo de producción, por lo que es necesario cambiar la configuración de la aplicación. Si los usuarios son desarrolladores, es decir que tienen rol de administrador, necesitaban acceder a AWS y realizar la configuración modificando el archivo JSON almacenado en S3. Al concluir con la configuración, iniciaron con el proceso de carga del archivo complejo. Para la carga del

archivo tuvieron que cambiar las ubicaciones de inicio de los datos acorde al diseño de la tabla de datos y realizar la descripción de los valores solicitados en el aplicativo.

Al culminar la evaluación el usuario quedó libre de probar la carga de archivos con formatos propios durante al menos 5 minutos, para luego llenar el Formulario de Retroalimentación.

3.3. Resultados

Resultados obtenidos en la etapa de Observación y de Retroalimentación. A continuación, se analizará cada etapa a profundidad.

3.3.1. Resultados etapa de observación

El plan de observación tiene el objetivo de recolectar datos sobre 5 pasos del proceso total de carga de un archivo en la aplicación. Puede visualizar todos los datos capturados en el Anexo 2.

El primer paso de la evaluación es cargar un archivo, por lo que como resultado de la prueba se obtuvieron los siguientes datos con respecto al tiempo que tardó el usuario para cumplir con la tarea.

Tiempo empleado por usuario para cargar un archivo

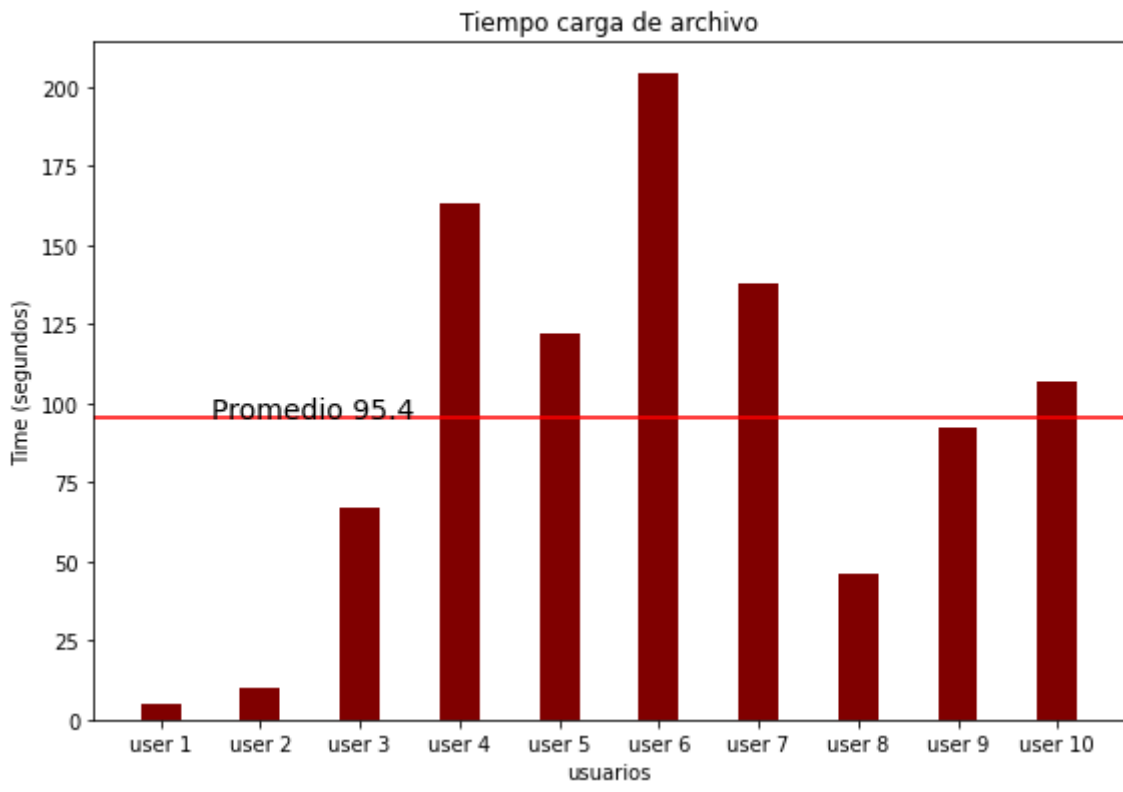


Ilustración 12 Tiempo empleado por usuario para cargar un archivo

En la Ilustración 12 se observan los tiempos que le tomó a cada usuario realizar la tarea de cargar un archivo donde el tiempo promedio fue de 95,4 segundos.

El segundo paso corresponde a la descripción del archivo de datos. En este paso 2 usuarios presentaron inconvenientes para comprender como realizar la descripción. Los usuarios pidieron ayuda debido a que no comprendían que es lo que tenían que hacer, generando así los siguientes tiempos para cumplir el proceso:

Tiempo empleado por usuario para describir un archivo

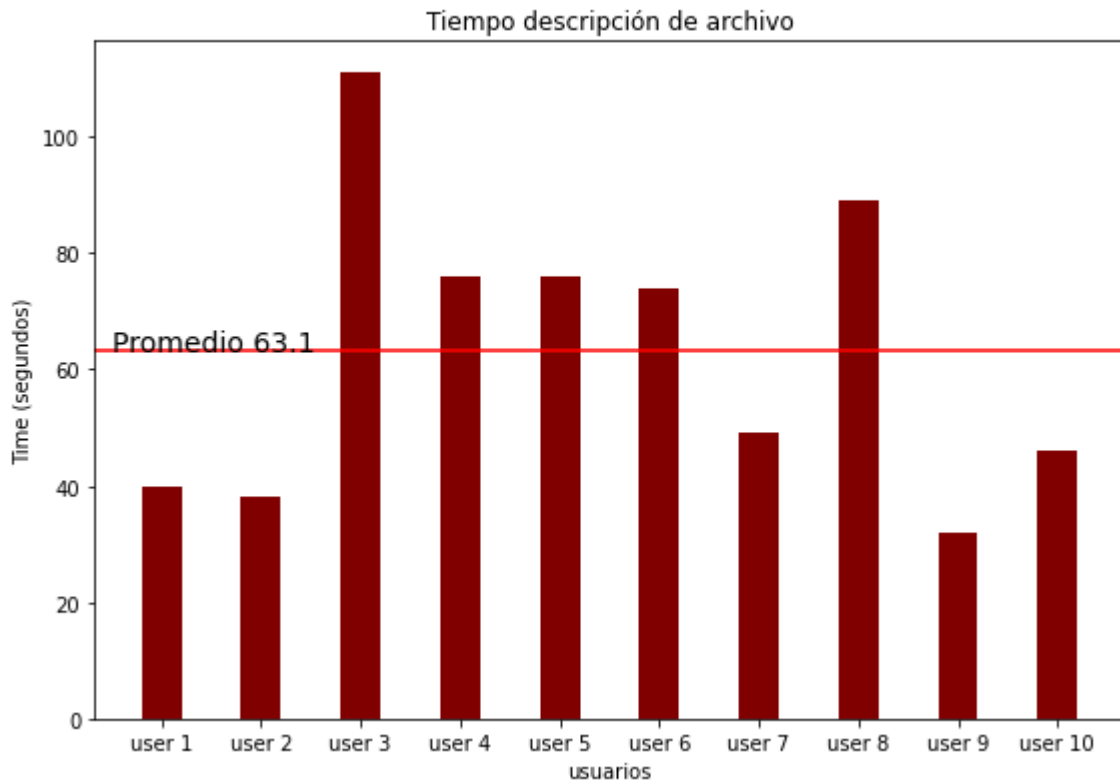


Ilustración 13 Tiempo empleado por usuario para describir un archivo

En la Ilustración 13 se visualizan los datos del tiempo empleado por los usuarios para describir un archivo de cálculo. Para el proceso de descripción se obtuvo un tiempo promedio de 1 minuto con 3 segundos.

En el tercer paso se analiza el tiempo que tarda el sistema en procesar el archivo para cada usuario, esta no es una métrica que califique al usuario, pero resulta útil para analizar el tiempo total empleado en el proceso completo de carga.

Tiempo que tarda el sistema en procesar un archivo por usuario

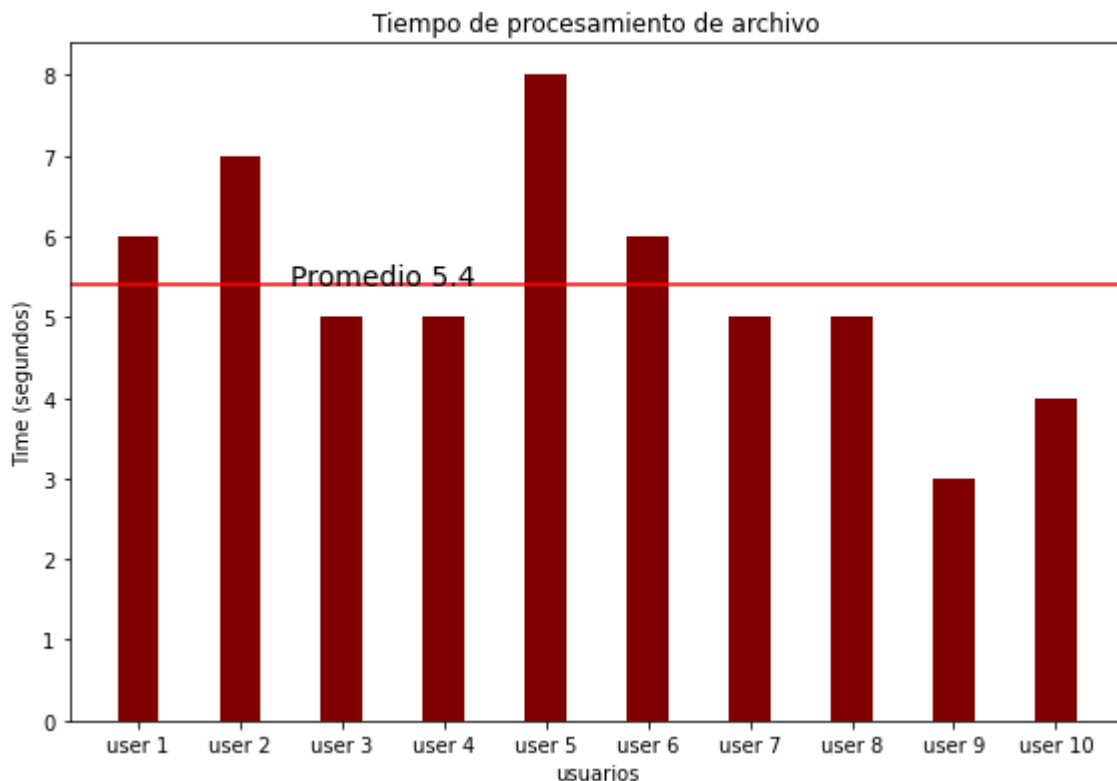


Ilustración 14 Tiempo que tarda el sistema en procesar un archivo por usuario

En la Ilustración 14 se observa el tiempo que le tomó al sistema procesar el archivo para cada uno de los usuarios. Se pudo observar que los tiempos más altos sucedían cuando se creaban por primera vez las locaciones en los buckets de S3, obteniendo un tiempo de procesamiento promedio de 5.4 segundos.

Una vez se procesa el archivo de datos inicia el siguiente paso que corresponde a encontrar el archivo cargado en la sección historial. Esto se analiza con el objetivo de determinar que tan intuitivo es el diseño del historial para los usuarios.

Tiempo que tardan los usuarios en encontrar el archivo que cargaron

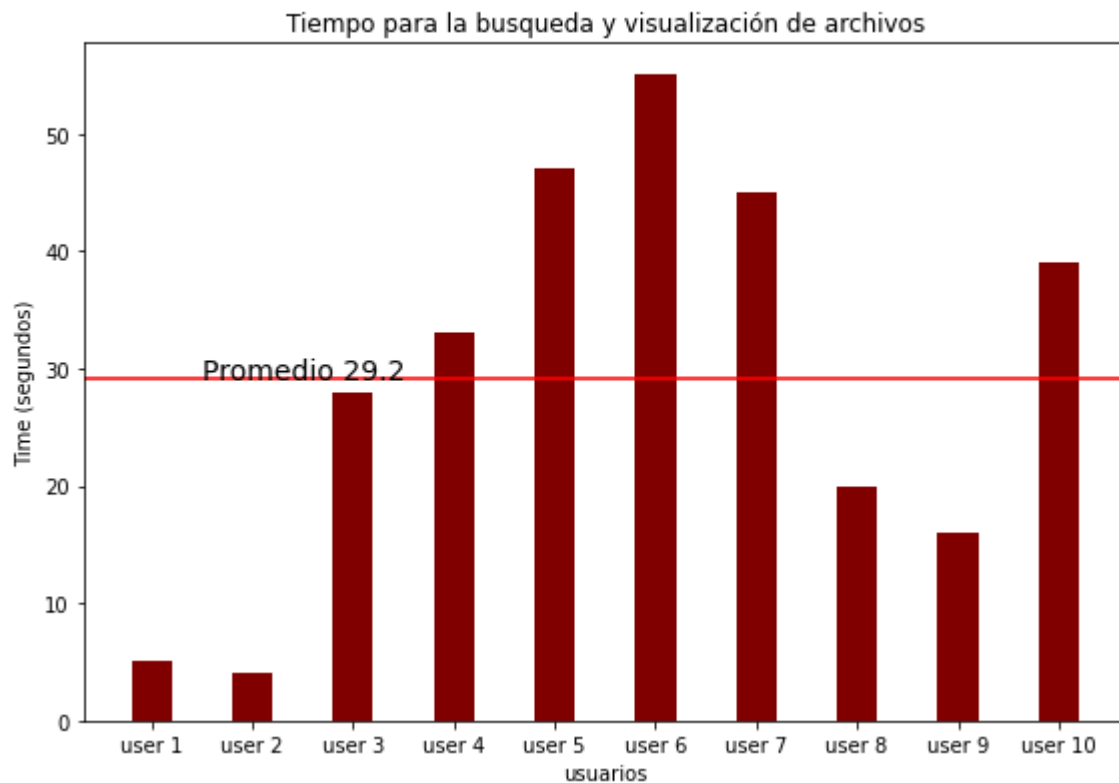


Ilustración 15 Tiempo que tardan los usuarios en encontrar el archivo que cargaron

Acorde al gráfico de la Ilustración 15, a los usuarios les toma en promedio 29.2 segundos encontrar los archivos procesados en la sección historial de la aplicación.

Por ultimo se evaluó el tiempo que toma el proceso de configuración de la aplicación.

Tiempo que tardan los usuarios en configurar la aplicación para que solicite nuevos valores

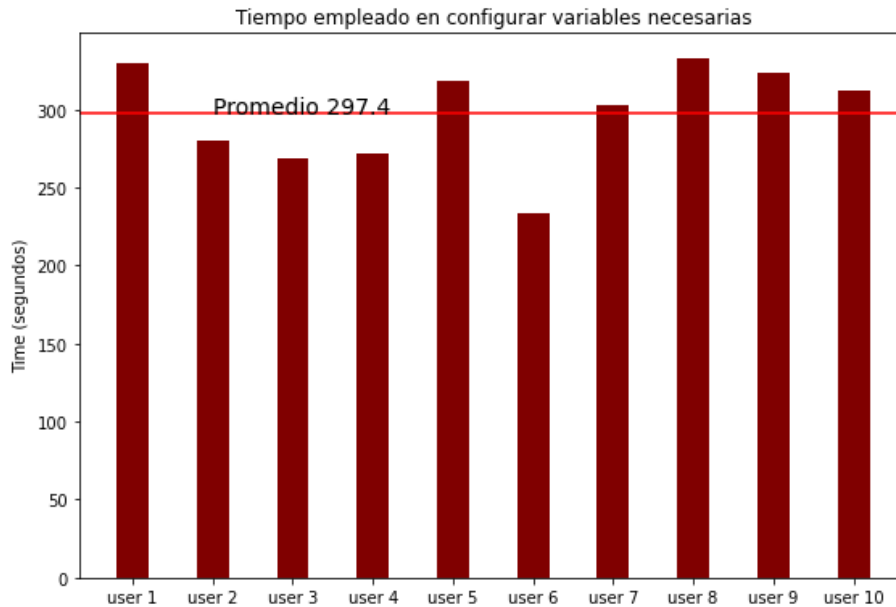


Ilustración 16 Tiempo que tardan los usuarios en configurar la aplicación para que solicite nuevos valores

En la Ilustración 16 se visualiza el tiempo que le tomó a cada usuario configurar un nuevo tipo de producción. Realizar la configuración les tomó a los usuarios en promedio 4 minutos con 57 segundos.

3.3.2. Resultados etapa de Retroalimentación

El formulario constaba de 8 preguntas cualitativas que permitieron recolectar las opiniones y comentarios de los usuarios evaluados. Puede visualizar el formulario completo en el Anexo 1.

La primera pregunta, valida el rol de los usuarios en Asimetrix para identificar las respuestas acordes a la relevancia que tiene la aplicación para el usuario.

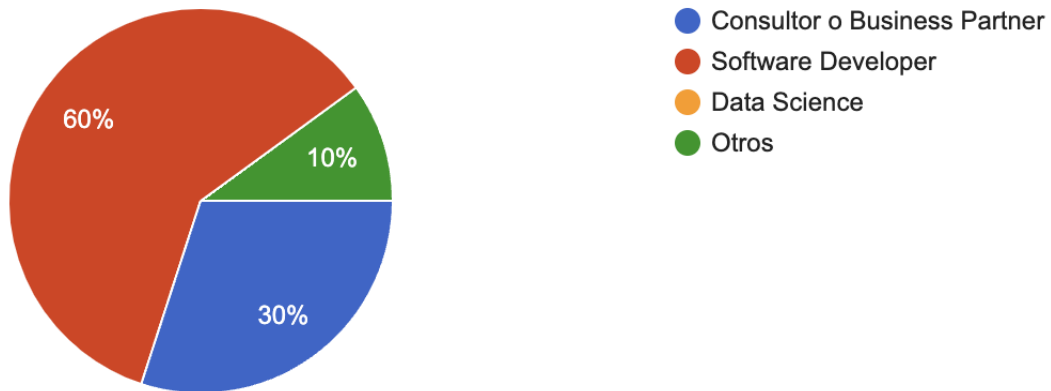


Ilustración 17 Roles de usuarios evaluados

Como se puede visualizar en el de la Ilustración 17 el 60% de los usuarios evaluados fueron administradores y el 30% consultores.

La segunda pregunta tenía como objetivo diferenciar las opiniones de los usuarios que interactuaban por primera vez con el aplicativo de los que formaron parte del proceso de validación de prototipos e historias de usuario iniciales.

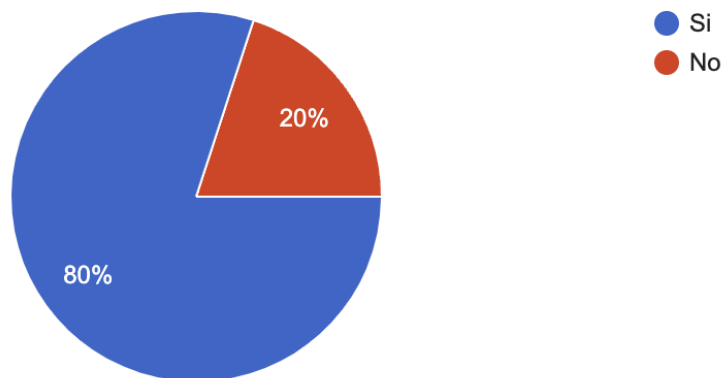


Ilustración 18 Usuarios que interactúan con la aplicación por primera vez

En la Ilustración 18 se observa que el 80% de los usuarios evaluados tuvieron interacción con la aplicación por primera vez durante el proceso de evaluación. El gráfico muestra que la evaluación se realizó mayormente con usuarios sin conocimiento previo del uso de la aplicación.

La tercera pregunta se solicitaba al usuario calificar con un número entre 1 y 5 la facilidad que tuvo al utilizar la aplicación. Donde 1 es Muy Difícil y 5 Muy Fácil.

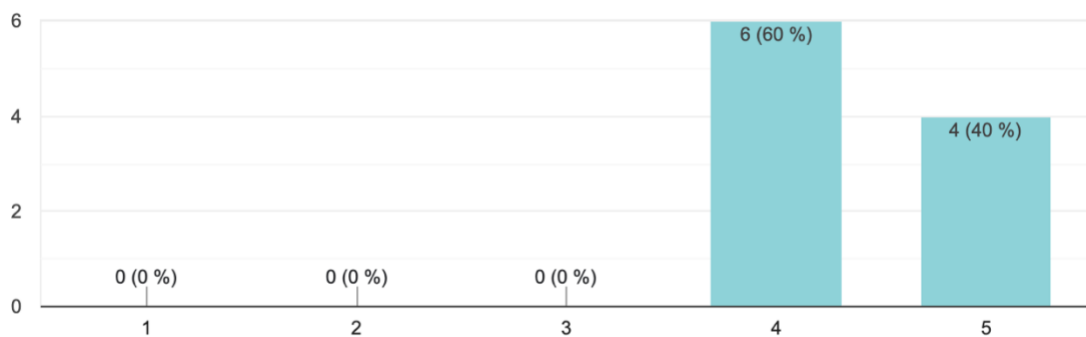


Ilustración 19 Facilidad de uso de la aplicación

En la Ilustración 19 se observa que todos los usuarios opinan que usar la aplicación es fácil, siendo para el 40% del total muy fácil de usar.

La cuarta pregunta mide los niveles de satisfacción considerando 3 aspectos cualitativos sobre la aplicación, que son: aspecto profesional, visualmente agradable y utilidad. Cada uno de los aspectos debía ser calificado indicando si cumple con las expectativas, supera las expectativas y por debajo de las expectativas.

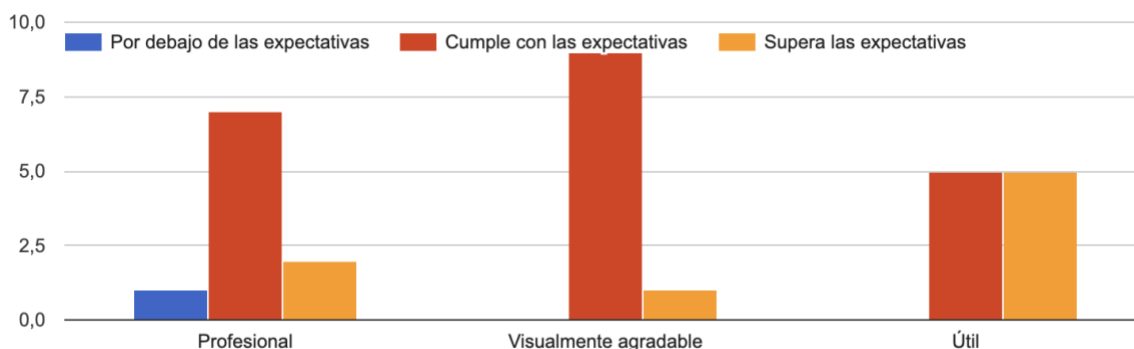


Ilustración 20 Expectativas sobre aspecto profesional, visual y utilidad

La Ilustración 20 permitió conocer que para los usuarios el aplicativo cumple con las expectativas.

En la quinta pregunta se consulta al usuario que es lo que más le gustó de la aplicación. En esta pregunta destacan las siguientes respuestas:

- “El historial de archivos procesados permitiendo descargar de manera simple tanto el archivo procesado como su versión original”
- “La posibilidad de procesar datos de manera simple en los excels.”
- “Me gustó mucho el concepto y la facilidad para generar mapeos de los datos”
- “Automatiza un proceso el cual se hace manual actualmente”
- “La rapidez, nos permitiría autonomía”

La sexta pregunta permitió recolectar las criticas constructivas y comentarios que aporten a mejorar el proyecto. Entre los comentarios más destacados están:

- “La forma de cargar los archivos es funcional, podría mejorar la experiencia de usuario implementar la opción de cargar múltiples archivos al tiempo”
- “El campo donde se indica la fila de inicio, inicia en un valor en 0, pero como el excel empieza en 1 puede llegar a no ser intuitivo que siempre se debe manejar el valor de la ubicación del excel menos 1.”

- “La casilla de selección de columnas debería desplazarse durante la visualización del archivo al momento de hacer scroll en cualquier dirección”
- “Agregar tooltips en los campos cuando se va a seleccionar en el momento de la carga.”

La séptima pregunta pedía al usuario comentar que inquietudes tuvo durante el proceso de evaluación. Entre las inquietudes más destacadas están:

- “Tuve dudas cuando procedí a indicar las columnas de donde se saca la información por variable, en un momento pensé que se iba a renombrar las columnas, pero luego me di cuenta de que se debe indicar la letra de la columna de donde se van a sacar los datos de cada una de las variables”
- “El render que muestra el archivo no tiene las mismas posiciones que el original entonces no sabía en que punto comenzaba la Fila inicial de los datos, me tuvieron que decir que era con respecto al original.”

La octava pregunta pidió a los usuarios brindar nuevas ideas que ayuden a mejorar el proyecto. Todos los usuarios brindaron muchas ideas útiles y valiosas entre las cuales destacan las siguientes:

- “Cargar varios archivos al tiempo”
- “Dentro de las extracciones tenemos datos que son fijos u otros que son calculados, creo que se debe tener en cuenta las formulas para los calculos y los datos fijos para la extracción”
- “Agregar información sobre las posibles variables. Ejemplo: Un texto explicando la variable, sus unidades, periodicidad, etc.”

3.4. Costos

Los costos de producción de la aplicación corresponden a los servicios de AWS que se necesitan para el despliegue y desarrollo de los microservicios.

Para el diseño de la se utilizaron servicios que están incluidos en la capa gratuita de AWS. Pero una vez se encuentre en producción la aplicación es muy probable que supere los límites mensuales bajo los que rige la gratuidad de los servicios. Por lo que se considerarían los siguientes precios estimados:

- **AWS S3**

Con almacenamiento mensual de 10 GB.

Almacenando archivos con un peso promedio de 50 MB.

1000 solicitudes PUT, POST

10000 solicitudes GET y SELECT

Retornando aproximadamente 50 GB de datos por mes

Total, por mes: 0.27 USD

- **AWS Amplify**

Con 5 minutos de proceso de despliegue por mes

2 GB de almacenamiento por mes

10 GB de datos servidos

2000 solicitudes SSR por hora

Duración de cada solicitud de 500 ms

Total, por mes: 42,60 USD

- **AWS Lambda**

Con arquitectura x86

300000 solicitudes por mes

3000 ms de ejecución por solicitud

500 MB de memoria asignada

1000 MB de almacenamiento efímero asignado

Total, por mes: 0.67 USD

- **AWS API Gateway**

Rest API

100000 solicitudes al mes

Sin cache

Total, por mes: 0.35 USD

AWS cobra por uso de sus servicios, por lo que los valores son solamente para brindar una noción de los costos que pueden llegarse a generar.

Con la estimación realizada se obtiene que al superar los límites de la capa gratuita el costo por mantener desplegada y funcionando la aplicación es de \$43.22 USD al mes.

Capítulo 4

4. Conclusiones y Recomendaciones

Este capítulo está compuesto por las conclusiones a las que se llegó tras el desarrollo de la aplicación y las recomendaciones a considerar para desarrollar de manera correcta el diseño de la aplicación.

4.1. Conclusiones

El desarrollo de la aplicación brinda grandes beneficios para Asimetrix y las empresas de analítica de datos para el procesamiento de archivos de datos. Entre los beneficios más destacados están:

- No es necesario que las empresas de producción de proteína animal deban estandarizar sus formatos de archivos de datos a los provistos por las empresas de analítica. Ya que con la interfaz pueden describir cualquier formato o diseño para generar un archivo ordenado acorde a lo solicitado por la empresa de analítica y procesarlo.
- El procesamiento de los archivos de datos es mucho más rápido usando la aplicación que con los procesos actuales de Asimetrix para realizar la misma tarea.
- Al ser el proceso de carga y procesamiento de archivos de datos muy fácil, tal como lo indican los usuarios en la retroalimentación, no sería necesario que un programador deba realizar el proceso. Reduciendo considerablemente los costos de procesamiento para Asimetrix.
- El tiempo que esperan las empresas pecuarias para que sus datos sean procesados con los métodos actuales de Asimetrix pueden tomar días o semanas, pero con el uso de la aplicación se realiza en cuestión de minutos.

El diseño y la infraestructura de la aplicación permite que se implementen nuevas funcionalidades con facilidad. Al estar desarrollada bajo el concepto de microservicios presenta ventajas como la escalabilidad y la funcionalidad modular.

4.2. Recomendaciones

- Para implementar el diseño de la aplicación se recomienda utilizar las tecnologías del diseño o buscar alternativas equivalentes.
- Tener en consideración que para el procesamiento de archivos de datos extremadamente grandes será necesario optimizar su renderizado, ya que puede generar lentitud por la alta exigencia de recursos que requiere. Para evitar dicha saturación se aconseja renderizar máximo 100 filas del archivo de datos y limitar la cantidad de caracteres que se muestran por celda de contenido.
- Es aconsejable desarrollar una interfaz de configuración sencilla ya que para el diseño actual es necesario tener conocimiento sobre el formato de texto JSON. Conocimiento que al no tenerse puede llegar a limitar los usuarios que puedan administrar la aplicación.
- Para reducir costos en producción podría considerarse cambiar el servicio de AWS Amplify por AWS Cloudfront el cual tiene mayor complejidad de uso, pero es mucho más económico.

Bibliografía

- datos, M. p. (2019). *Ideca*. Recuperado el octubre de 2022, de https://www.ideca.gov.co/sites/default/files/MetodologiaAnaliticaDatos_0.pdf
- El negocio de los datos, e. i. (2020). *Comillas, Universidad Pontificia de Madrid*. Obtenido de [repositorio.comillas: https://repositorio.comillas.edu/xmlui/bitstream/handle/11531/42729/EI%20Negocio%20de%20los%20Datos.%20EI%20Impacto%20del%20Business%20Analytics%20en%20la%20Empresa%20-%20201600314.pdf?sequence=-1&isAllowed=y](https://repositorio.comillas.edu/xmlui/bitstream/handle/11531/42729/EI%20Negocio%20de%20los%20Datos.%20EI%20Impacto%20del%20Business%20Analytics%20en%20la%20Empresa%20-%20201600314.pdf?sequence=-1&isAllowed=y)
- datos, A. d. (2019). *Cámara de comercio de Cali*. Obtenido de <https://www.ccc.org.co/inc/uploads/2019/05/Guia-analitica-de-datos.pdf>
- Big Data: una exploración de investigaciones, t. y. (2017). *Scielo*. Obtenido de <http://www.scielo.org.co/pdf/teclo/v20n39/v20n39a02.pdf>
- Occidental, L. I. (2021). *IICA*. Obtenido de <https://repositorio.iica.int/bitstream/handle/11324/16954/BVE21068221e.pdf?sequence=1&isAllowed=y>
- producción, L. e. (s.f.). *mheducation*. Obtenido de <https://www.mheducation.es/bcv/guide/capitulo/8448149971.pdf>
- metodos, E. d. (2015). Obtenido de Yusef Hassan Montero: https://yusef.es/Experiencia_de_Usuario.pdf
- competitividad, A. i. (2020). *FACPYA*. Obtenido de http://www.web.facpya.uanl.mx/vinculategica/Vinculategica6_2/37_Rosales_Arecha_vala.pdf
- organizaciones, L. a. (2020). *FACPYA*. Obtenido de http://www.web.facpya.uanl.mx/Vinculategica/Vinculategica6_2/5_Trevi%C3%B1o_Rivera_Garza.pdf
- world, T. a.-d. (2016). *cosmeticinnovation*. Obtenido de <https://www.cosmeticinnovation.com.br/wp-content/uploads/2017/01/MGI-The-Age-of-Analytics-Full-report.pdf>

Anexos

Anexo 1

FORMULARIO DE RETROALIMENTACIÓN

¿Cual es su rol en la compañía ? *

- Seleccione el rol con el que más sienta afinidad.

Si no tiene afinidad con ninguno seleccione: Otros

Consultor o Business Partner

Software Developer

Data Science

Otros

¿Es la primera vez que visualiza el proyecto? *

Si

No

¿Qué tan facil fue realizar el proceso de carga de un archivo? *

1 2 3 4 5

Muy Dificil Muy fácil

Ilustración 21 Formulario Google parte 1

¿Cuál es su impresión general de la aplicación? *

	Por debajo de las expectativas	Cumple con las expectativas	Supera las expectativas
Profesional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Visualmente agradable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Útil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Qué es lo que más le gustó de la aplicación?

Tu respuesta _____

Comentarios que aporten a mejorar el proyecto.
(sobre lo ya implementado)
- Críticas constructivas

Tu respuesta _____

Ilustración 22 Formulario Google parte 2

¿Que inquietudes tuvo durante el proceso de evaluación?

Tu respuesta _____

Ideas nuevas que ayuden a mejorar el proyecto

Tu respuesta _____

Ilustración 23 Formulario Google parte 3

Anexo 2

Datos Capturados Plan de Observación

Objetivos/Personas										
Carga de Archivo	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
Tiempo promedio	0:05	0:10	1:07	2:43	2:02	3:24	2:18	0:46	1:32	1:47
Se realizó con éxito	si	si	si	si	si	si	si	si	si	si
complejidad	baja	baja	baja	baja	media	media	bajo	media	baja	baja
Intentos	1	1	1	1	1	1	1	1	1	1
Descripción de archivo	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
Tiempo promedio	0:40	0:38	1:51	1:16	1:16	1:14	0:49	1:29	0:32	0:46
Se realizó con éxito	si	si	si	si	si	si	si	si	si	si
Intentos	1	1	1	1	1	2	1	1	1	1
# Errores generados	0	0		0	0	0	0	0	0	0
Busqueda y visualización de archivos	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
Tiempo promedio	0:05	0:04	0:28	0:33	0:47	0:55	0:45	0:20	0:16	0:39
Se realizó con éxito	si	si	si	si	si	si	si	si	si	si
Intentos	1	1	1	1	1	1	1	1	1	1
# Errores generados	1	1	0	0		0	0	0	0	0
Configuración variables	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
Tiempo promedio	5:30	4:40	4:29	4:32	5:18	4:14	5:03	5:33	5:23	5:12
Se realizó con éxito	si	si	si	si	si	si	si	si	si	si
Intentos	1	1	1	1	1	1	1	1	1	1
# Errores generados	0	0	0	0	0	0	0	0	0	0
Procesamiento	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
# Errores generados	0	0	0	0	0	0	0	0	0	0
Tiempo promedio	0:06	0:07	0:05	0:05	0:08	0:06	0:05	0:05	0:03	0:04

Ilustración 24 Datos capturados plan de observación

