

**“PROPUESTA PARA DETERMINAR LA CANTIDAD Y PREFERENCIAS DE USO QUE LE DAN
LOS SUSCRIPTORES A LAS HERRAMIENTAS Y APLICACIONES DATASOL”**

AUTOR:

JESUS ALBERTO MARTINEZ - CÓDIGO 1722010341

GRUPO 3

ASESOR: MSC. GIOVANNY ALEXANDER BAQUERO VILLAMIL

INSTITUCIÓN UNIVERSITARIA POLITÉCNICO GRANCOLOMBIANO

FACULTAD DE INGENIERÍA Y CIENCIAS BÁSICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ESPECIALIZACIÓN EN GERENCIA DE PROYECTOS EN INTELIGENCIA DE NEGOCIOS

BOGOTÁ, D.C. 2018

Tabla de contenido

1. Título	4
2. Resumen.....	4
2.1. Español	4
2.2. Inglés	4
3. Tema.....	5
4. Dedicación.....	6
5. Fundamentación del proyecto	6
5.1. Marco contextual	6
5.1.1. Empresa.....	6
5.1.2. Misión.....	7
5.1.3. Visión.....	7
6. Problema	7
7. Justificación.....	8
7.1. Objetivo General	9
8. Marco conceptual	9
9. Estado del arte	11
9.1. Estudio sobre la experiencia visual de la interfaz del sitio web de compras para personas mayores.....	11
9.2. El análisis de rendimiento de los sistemas de almacenamiento distribuido utilizados en sistemas web escalables.	11
9.3. Predicción de la población basada en los modelos múltiples y la investigación comparativa.	12
9.4. Detección de amenazas internas basadas en escenarios a partir de actividades cibernéticas	12
9.5. Un modelo para el diseño inclusivo de bibliotecas digitales.	12
9.6. LabBook: análisis de datos de colaboración social impulsado por metadatos	13
9.7. Hacia el desarrollo de un marco de predicción de errores web basado en series temporales convencionales.....	13
9.8. Caracterizar el interés del usuario en las bases de datos NoSQL de datos de preguntas y respuestas sociales.....	14
9.9. Sistema de plataforma de computación de comercio electrónico basado en información de logística distribuida	14

9.10. La importancia relativa de los factores de usabilidad y funcionalidad para los sitios web de ciber salud.	15
9.11. ¿Qué es la usabilidad en el contexto de la biblioteca digital y cómo puede medirse?	16
10. Objetivos Específicos, actividades y cronograma	16
10.1. Objetivos Específicos.....	16
10.2. Metodología	17
10.3. Presupuesto general del proyecto	18
11. Viabilidad financiera.....	19
12. Plan de actividades.....	19
13. Plan de adquisiciones	21
14. Plan de riesgos	21
15. Plan de interesados	22
16. Conclusiones y recomendaciones	23
17. Referencias.....	24

1. Título

Propuesta para determinar la cantidad y preferencias de uso que le dan los suscriptores a las herramientas y aplicaciones DATASOL para enfocar los contenidos, utilitarios y estrategias comerciales con las necesidades de los clientes y así generar una mejor satisfacción y minimizar su deserción.

2. Resumen

2.1. Español

Para alinearse organizacionalmente y cumplir con la misión y visión establecidas se requiere identificar los patrones de los clientes tanto en el aspecto de hábitos de compra como en el tipo de contenidos de su interés, segmentándolos adecuadamente para que la información que se les provea sea primordial y de gran aporte en su gestión.

Nada de esto será posible si las unidades estratégicas y de apoyo de la empresa no unen esfuerzos para cumplir con un fin común y no en resultados particulares. Cada área tiene un conocimiento e información parcial que solo muestra una vista del cliente y que mediante una adecuada consolidación de los datos y una manipulación responsable se va a vislumbrar una aproximación más real de los clientes en donde será más natural proveerles unos productos y servicios adecuados a sus necesidades, generando nuevas herramientas que complementen y enriquezcan la gestión diaria de los usuarios.

No se descarta la posibilidad del surgimiento de nuevos negocios o modelos de negocio a partir de la información recopilada la cual va a mostrar nuevas dimensiones que hasta hoy no se habían evidenciado de los clientes DATASOL.

2.2. Inglés

In order to align organizationally and comply with the established mission and vision, it is necessary to identify the clients' patterns both in the aspect of purchasing habits and in the type

of content of interest, segmenting them appropriately so that the information provided to them is essential and the great contribution in its management.

Nothing will be possible if the strategic and support units of the company do not join efforts to fulfill a common goal and not in particular results. Each area has a knowledge and partial information that only shows a view of the client and that through an adequate consolidation of the data and a responsible manipulation is going to glimpse a more real approach of the clients where it will be more natural to provide them with adequate products and services to their needs, generating new tools that complement and enrich the daily management of users.

The possibility of the emergence of new businesses or business models from the information collected is not ruled out, which will show new dimensions that until now had not been evidenced by the DATASOL clients.

3. Tema

Esta propuesta enmarca los elementos y reglas a tener en cuenta para determinar de manera unificada y consolidada por cliente la adquisición y uso que éste da a las herramientas y contenidos DATASOL permitiendo al equipo de trabajo de esta organización sugerir mejoras o cambios a sus productos enfocándolas en las necesidades, tendencias e inquietudes de sus usuarios.

Al tener mayor conocimiento del accionar de sus clientes le dará a la organización una perspectiva que la acerca más a sus suscriptores obteniendo de primera mano comportamientos y enfoques por segmentos, tamaño y tipo de organización.

También podrá identificar patrones para evitar malas negociaciones a futuro ofreciendo una mejor propuesta comercial a sus clientes según el perfil que tengan.

Así mismo, enfocar esfuerzos invirtiendo en módulos en donde la penetración del mercado sea mas atractiva y dejando de lado herramientas y contenidos de poco interés.

Todo esto redundará en un mejor servicio al cliente atendiéndolo de manera integral dependiendo de los productos que haya adquirido y del perfilamiento del mismo aumentando la retención de clientes y un mejor nivel de satisfacción.

4. Dedicación

Tipo de Actividad	Sub-actividad	% de Dedicación
Investigación teórica	N/A	20%
Obtención de datos	Contacto telefónico	10%
	Estadísticas de uso	15%
	Contacto electrónico (Correo/Quejas y Reclamos)	10%
	Listados de clientes vigentes y no vigentes	5%
Diseño del Proyecto	N/A	40%

5. Fundamentación del proyecto

5.1. Marco contextual

5.1.1. Empresa

DATASOL nace como respuesta ante la ausencia de una entidad estatal que divulgara las leyes de manera permanente y en forma comprensible y sencilla, para que tanto el abogado como el empresario y el mismo estado, contaran con información legal organizada, completa y actualizada periódicamente.

Así se lanza al mercado **el primer ejemplar de la revista Legislación Económica**, publicación que compila quincenalmente las disposiciones legales colombianas del orden nacional en todas las áreas de interés general.

De la experiencia de los fundadores en el ejercicio de la profesión de abogados, nace la idea de realizar unas compilaciones jurídicas que fuesen una herramienta de trabajo diario. La idea provino de la necesidad de tener a la mano, en un solo tomo actualizable, lo esencial de cada disciplina, perfeccionando una metodología compilativa que permitiera encontrar en un solo sitio la información pertinente a la materia que se esté consultando: las normas principales, complementarias, concordancias, extractos jurisprudenciales y doctrinales, ejemplos de aplicación práctica y minutas y por eso se inicia la entrega en producto físico de compilaciones legales por campos del derecho público, privado y de materias especializadas.

Con la experiencia y apoyada en su exitosa metodología editorial, DATASOL se convirtió en una Multinacional colombiana con presencia en Argentina, México, Chile, Perú, y Venezuela, con un total de 1.250 empleados.

En la sede principal, se encuentra organizada por:

Unidades Estratégicas de Negocio (UENs): Información Jurídica, Información Profesional Especializada, Oficina, Escolar y Promocional e Impresos.

Unidades de Apoyo (UEAs): que soportan las actividades de Finanzas, Jurídica, Logística, Tecnología, Innovación y Desarrollo y Gestión Humana.

5.1.2. Misión

En DATASOL proveemos soluciones efectivas de información innovadora de tal forma que nuestros clientes puedan tomar decisiones seguras sobre situaciones legales, aplicar para un empleo o contratar a alguien, investigar y profundizar en temas específicos, utilizar herramientas de uso profesional para acceder a conocimiento experto, construir un inmueble, hacer inteligencia de mercado y comunicarse con un cliente potencial u organizar un archivo.

5.1.3. Visión

Para el 2018, consolidarnos como el mejor proveedor de soluciones en información efectiva, posicionándonos en la mente de nuestros clientes como la primera opción para conservar, organizar, archivar, actualizar e imprimir información operando de forma sinérgica con nuestras unidades estratégicas de negocio trabajando en la misma dirección con nuestros clientes y usuarios.

6. Problema

Algunos clientes no aprovechan en su totalidad los productos DATASOL lo cual genera cambios decrecientes en los planes de suscripción o en el peor de los casos la deserción de los mismos. Al mes están desertando un 2,3% de los clientes y un 10% reduce su volumen de suscripción, que, aunque en cifras es menor que los clientes nuevos o renovados, reducen el margen de ganancia.

Se deben identificar los contenidos y herramientas tanto exitosos como menos exitosos para fortalecer los primeros y reestructurar o descartar los segundos acoplándose a las necesidades de los usuarios DATASOL.

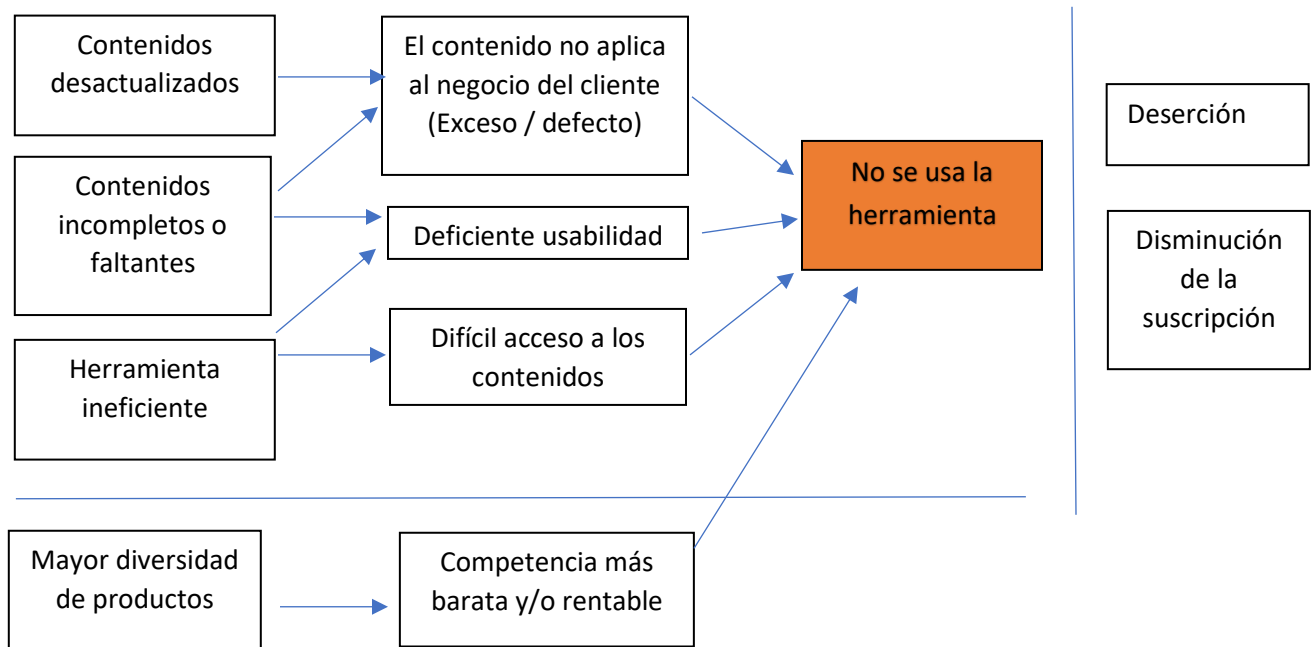


Figura 1. Planteamiento del problema

7. Justificación

En el mundo digital, las suscripciones a los servicios y productos que se ofrecen es la razón de ser de los mismos, más cuando estas son pagas representando ingresos y ganancias monetarias a una organización.

Del aumento de dichas suscripciones y de poder mantener un costo rentable y competitivo depende la existencia de un producto en una organización.

Pero dichos productos deben evolucionar y deben ser evaluados, para determinar cómo sus contenidos o funcionalidades van a proporcionar un beneficio al usuario que los utilice. Un documento o funcionalidad exitosa hoy, no implica que sea exitosa siempre. En ocasiones, algunos elementos conforman un producto y nunca son explorados por los usuarios tienden a desaparecer. La actualización de la información y de la funcionalidad es una constante.

Tener la capacidad de identificar cuales elementos que conforman un producto es de interés o no, es usado o no, hace la diferencia para poder enfocarlo y matizarlo a las necesidades del

mundo de hoy, el cual siempre es cambiante y en donde la competencia como demás fuentes de información y herramientas tratan de sobresalir ofreciendo más y mejores opciones para sus productos a precios muy competitivos.

Realizando el seguimiento a la gestión de los usuarios dentro de los productos DATASOL permitirá tener un punto de referencia objetivo sobre el cual basar parte de la evolución de los mismos.

7.1. Objetivo General

Establecer los elementos, reglas y procedimientos para determinar el nivel de satisfacción de un cliente y la razón por la cual permanece, modifica o cancela una suscripción con DATASOL tomando como base los datos de encuestas de satisfacción, quejas, reclamos, inquietudes, de acceso y de uso a los contenidos y herramientas y los datos comerciales históricos, identificando patrones o tendencias para generar o mejorar los productos y para perfilar adecuadamente al cliente y realizar una oferta acorde a sus necesidades reduciendo así la tasa de deserción y aumentando las ganancias de la organización.

8. Marco conceptual

En el contexto mundial, el incremento de la interacción comercial ha producido cambios importantes en el modo de competir. Esto ocasiona el cambio a un mundo sin fronteras, que obliga a las organizaciones a implementar prácticas que conduzcan a la adaptación a estas nuevas condiciones. De la misma manera, los países en las diferentes regiones adoptan estrategias, que les permitan favorecer su base productiva ampliando los mercados.

Es conveniente impulsar la innovación en gestión resaltando la necesidad e importancia de introducir nuevas prácticas y herramientas de gestión en nuestras organizaciones.

En la organización cada unidad de negocio o de apoyo cuenta con sus propios informes de comportamiento de los clientes, pero estos solo son parte de la visión global que deben tener sobre estos y sus preferencias.

Por ejemplo, el área comercial o financiera tiene información de cuantos clientes desertaron, renovaron o son nuevos. Pero el uso del servicio en sí, que módulos son de mayor interés lo sabe con mayor certeza el área de contenido. Ahora, si se analiza en cuales módulos han tenido

mayores inconvenientes o mayores dudas, el área de servicio al cliente es la que tiene dicha información.

Es muy factible que un reporte en ventas sea muy bueno en cifras, o por lo menos reporte números “en verde”, pero el uso o no uso de las aplicaciones y el no cuidar al mayor insumo de la compañía, el cliente, hace que día tras día se reduzca el mercado para los productos de la organización.

Al no tener esta información consolidada y centralizada, hace que las diferentes unidades requieran de informes y análisis que les permitan ver a sus clientes de otra manera, no parcialmente sino teniendo en cuenta más aspectos.

Las áreas interesadas en dicha información son:

- Unidades de negocio
 - o Gerencia
 - o Líderes de producto
- Mercadeo
- Financiera
- Comercial
 - o Gerencia
 - o Líder de cuenta
- Tecnología
 - o Infraestructura
 - o Líder técnico de producto
- Servicio al cliente

Los datos que se manejarán inicialmente empezarán a vislumbrar información que, en comparación con la inicialmente manejada, será más completa y cercana a la realidad. Este proceso se tendrá que complementar cada vez más mejorando los sistemas de captura de información para que la consolidación, depuración y normalización de la misma garanticen mejores resultados con adecuados tiempos de entrega y fiabilidad.

9. Estado del arte

9.1. Estudio sobre la experiencia visual de la interfaz del sitio web de compras para personas mayores.

Basado en la experiencia visual de los adultos mayores, este estudio analiza la relación entre los efectos visuales, los efectos de interacción y los efectos de experiencia producidos por las interfaces de sitios web de compras que están organizados por diferentes diseños y diferentes colores. Resumió la teoría del sistema de diseño de sitios de compras con la ayuda de experimentos relevantes, cuestionarios y estadísticas de datos, este estudio combina la psicología cognitiva de las personas mayores con la ergonomía y otras disciplinas para organizar sistemáticamente los principios de diseño de los sitios de compras y concluir un conjunto de interfaces de sitios web. *(Study on the visual experience of senior people shopping website interface, 2018, Julio 25)*

9.2. El análisis de rendimiento de los sistemas de almacenamiento distribuido utilizados en sistemas web escalables.

Los sistemas web escalables están directamente relacionados con los sistemas de almacenamiento distribuido que se utilizan para procesar grandes cantidades de datos (big data). Un ejemplo de un sistema de este tipo es Hadoop, con sus numerosas extensiones que admiten el almacenamiento de datos, como los sistemas SQL-on-Hadoop y el formato de archivo "Parquet". Otro tipo de sistemas para almacenar y procesar big data son las bases de datos NoSQL, como HBase, que se utilizan en aplicaciones que requieren un acceso aleatorio rápido. El sistema Kudu fue creado para combinar las ventajas de Hadoop y HBase y permitir tanto un análisis efectivo de conjuntos de datos como un acceso aleatorio rápido. Como objeto de la investigación, se realizó el análisis de rendimiento de los sistemas mencionados. El experimento se realizó en el entorno cloud donde se configuró el clúster de nueve máquinas virtuales. Para fines de investigación, se utilizó un fragmento de aproximadamente mil millones de filas del conjunto de datos público. Los resultados de las mediciones confirman que el sistema Kudu es una alternativa prometedora a las tecnologías de uso común. *(The performance analysis of distributed storage systems used in scalable web systems, 2018, Septiembre 18)*

9.3. Predicción de la población basada en los modelos múltiples y la investigación comparativa.

Primero, se establecen cuatro modelos para ajustar los datos de la población para el año 2000-2016. Los modelos incluyen el modelo GM (1, 1), el modelo de regresión lineal unaria, el modelo de índice y el modelo de crecimiento logístico. En segundo lugar, se realiza una simulación para demostrar los cuatro modelos. Los resultados de los datos del anuario estadístico y el sitio web muestran que el efecto de ajuste es bueno, y con alta precisión, por lo tanto, utilizamos cuatro modelos para predecir la población de 2017-2020. Luego, se construye un modelo de predicción combinado derivado de los cuatro modelos, que es más preciso que el modelo de predicción único indicado a partir de los resultados de ajuste. (*Population prediction based on the multi-models and comparison research, 2017, Diciembre 10*)

9.4. Detección de amenazas internas basadas en escenarios a partir de actividades cibernéticas

Un escenario de amenaza interno se refiere al resultado de un conjunto de actividades maliciosas causadas por el uso indebido intencional o no intencional de los sistemas, redes, datos y recursos de la organización. La prevención de amenazas internas es difícil, ya que participan personas de confianza de la organización que tienen acceso autorizado a estos recursos confidenciales. La investigación sobre detección de amenazas internas se centra principalmente en el desarrollo de técnicas de detección de anomalías de comportamiento no supervisadas con el objetivo de descubrir anomalías o cambios anormales en el comportamiento del usuario a lo largo del tiempo. Sin embargo, una actividad anómala no es necesariamente maliciosa, lo que puede llevar a una situación de amenaza interna. (*Scenario-based insider threat detection from cyber activities, 2018, Septiembre 3*)

9.5. Un modelo para el diseño inclusivo de bibliotecas digitales.

Las bibliotecas digitales (DL) deben atender no solo a las diversas necesidades de sus usuarios objetivo, sino también a sus diferentes capacidades y a las tecnologías de adaptación utilizadas por las personas cuyas capacidades informáticas están restringidas debido a discapacidades. Este

documento propone un modelo para el diseño de DL que incluye la optimización de la usabilidad del proceso de búsqueda y asegura la accesibilidad del contenido para los usuarios de DL con discapacidades. (*A model for inclusive design of digital libraries, 2007, Junio 23*)

9.6. LabBook: análisis de datos de colaboración social impulsado por metadatos

Se están adoptando plataformas abiertas de análisis de datos para apoyar la colaboración en ciencia y negocios. Los estudios sugieren que el trabajo analítico en una empresa ocurre en un complejo ecosistema de personas, datos y software que trabajan de manera coordinada.

Estos estudios también apuntan a la fricción entre los elementos de este ecosistema que reduce la productividad del usuario y la calidad del trabajo. LabBook es una plataforma de análisis de datos abierta, social y colaborativa, diseñada explícitamente para reducir esta fricción y acelerar el descubrimiento. Su objetivo es ayudar a los usuarios a aprovechar el conocimiento y la experiencia de cada uno para encontrar los datos, las herramientas y los colaboradores que necesitan para integrar, visualizar y analizar los datos. La idea clave es recopilar y utilizar más metadatos sobre todos los elementos del ecosistema analítico mediante una arquitectura y experiencia de usuario que reduzcan el costo de aportar dichos metadatos. (*LabBook: Metadata-driven social collaborative data analysis, 2015, Diciembre 22*)

9.7. Hacia el desarrollo de un marco de predicción de errores web basado en series temporales convencionales

La confiabilidad de la web está ganando importancia con el tiempo debido al aumento exponencial en la popularidad de diferentes redes sociales comunitarias, sistemas de correo y otras aplicaciones en línea. Por lo tanto, para mejorar la confiabilidad de cualquier sistema web existente, los administradores web deben tener conocimiento de varios errores web presentes en el sistema, influencias de varias características de carga de trabajo en la manifestación de varios errores web y las relaciones entre diferentes características de carga de trabajo.

Pero en realidad, a menudo puede que no sea posible establecer una correspondencia generalizada entre varias características de la carga de trabajo. Además, los problemas como la predicción y la estimación de las ocurrencias acumulativas de las fallas de contenido de origen y el tiempo

correspondiente entre fallas de un sistema web se vuelven menos destacados por la comunidad de investigación de confiabilidad. Por lo tanto, en este trabajo, los autores han presentado un procedimiento bien definido (un marco de pronóstico) para que los administradores de la web analicen y aumenten la confiabilidad de los sitios web bajo su supervisión. Inicialmente, toma el acceso HTTP y los registros de errores para extraer toda la información necesaria relacionada con las cargas de trabajo, los errores de la web y el tiempo correspondiente entre fallas. (*Toward the development of a conventional time series-based web error forecasting framework, 2018, Abril 1*)

9.8. Caracterizar el interés del usuario en las bases de datos NoSQL de datos de preguntas y respuestas sociales.

Con el advenimiento de la tecnología de redes sociales para compartir preguntas y respuestas frecuentes entre los usuarios finales, hay un interés cada vez mayor en comprender las características y en utilizar datos de preguntas y respuestas sociales (QA). No solo SQL (NoSQL) es un tema técnico popular en los sitios web que responden preguntas sociales y está ganando popularidad con las nuevas demandas de bases de datos escalables de big data. A pesar del gran interés de los usuarios en la tecnología NoSQL, aún no se ha realizado un intento de analizar cómo reaccionan los usuarios reales ante NoSQL. Por lo tanto, en el presente trabajo, utilizamos los datos de control de calidad adquiridos de Stack Overflow (un sitio web de control de calidad que funciona como un gran repositorio de conocimientos) para comprender cómo las personas perciben la tecnología NoSQL. Con este fin, se utilizan técnicas de modelado de asignaciones de Dirichlet latentes (LDA) para descubrir la tendencia de las bases de datos NoSQL. Además, examinamos un modelo LDA ponderado para reflejar la influencia de las respuestas y, finalmente, proponemos el valor de discriminación del tema en un intento de encontrar temas que distingan cada base de datos NoSQL. (*Characterizing user interest in NoSQL databases of social question and answer data, 2018, Marzo 7*)

9.9. Sistema de plataforma de computación de comercio electrónico basado en información de logística distribuida

Los sitios web de comercio electrónico generan una gran cantidad de datos de comportamiento de los usuarios, con el aumento continuo del volumen de negocios de las empresas de comercio electrónico. Las empresas esperan tener un profundo conocimiento de cada cliente a través de estos datos y esperan formar una relación de aprendizaje con los clientes. En base a esto, este documento

elabora primero el trasfondo y la importancia del desarrollo del sistema, el status quo de las tecnologías relacionadas en el país y en el extranjero, luego lleva a cabo el diseño general y ofrece el plan de realización del sistema, y finalmente diseña y realiza el big data todo esto basado en el marco de computación en tiempo real distribuido, las Empresas, especialmente las compañías de Internet, pueden proporcionar soporte de datos interno mediante la construcción de una infraestructura de almacenamiento de datos jerárquica y tecnología de computación de datos en tiempo real, y luego agrupar, dividir y predecir datos más relevantes a través de estadísticas matemáticas y algoritmos de minería.

Cada uno de nosotros está etiquetado con descripciones para obtener cada uno de nuestros atributos y pasatiempos, etc., que brindan un sólido apoyo a las necesidades externas de las empresas y al rápido negocio basado en datos. *(E-commerce big data computing platform system based on distributed computing logistics information, 2018, Febrero 23)*

9.10. La importancia relativa de los factores de usabilidad y funcionalidad para los sitios web de ciber salud.

Con el uso cada vez mayor de Internet, la demanda de información y asesoramiento sobre atención médica en línea, así como la cantidad de sitios web relacionados con la salud, están aumentando. En caso de información y consejos de salud en línea, la interfaz de usuario reemplaza la comunicación cara a cara. Para garantizar que se satisfagan las necesidades de los usuarios, es fundamental equilibrar la funcionalidad y la facilidad de uso en el diseño del sitio web.

El presente estudio busca identificar las complejas interrelaciones entre los diversos factores de usabilidad y funcionalidad en relación con los sitios web de salud electrónica. Se evaluaron dos sitios web turcos de salud electrónica para su evaluación en este estudio. Los hallazgos muestran que los usuarios de los sitios web de información de salud otorgan una mayor prioridad a la funcionalidad y sus factores, por lo que la mayor importancia relativa está en los "servicios/instalaciones" y la "personalización/categorización de la información". Los factores de usabilidad más importantes relacionados con los sitios web de e-health son la "memorización" y la "interacción". *(The relative importance of usability and functionality factors for E-health web sites, 2013, Julio 4)*

9.11. ¿Qué es la usabilidad en el contexto de la biblioteca digital y cómo puede medirse?

Este documento revisa cómo se ha definido la usabilidad en el contexto de la biblioteca digital, qué métodos se han aplicado y su aplicabilidad, y propone un modelo de evaluación y un conjunto de instrumentos para evaluar la usabilidad para bibliotecas digitales académicas. El modelo examina la efectividad, la eficiencia, la satisfacción y la capacidad de aprendizaje. Se ha encontrado que existe una relación entrelazada entre eficacia, eficiencia y satisfacción. También examina cómo la capacidad de aprendizaje interactúa con estos tres atributos. *(What is usability in the context of the digital library and how can it be measured?, 2005, Junio)*

10. Objetivos Específicos, actividades y cronograma

10.1. Objetivos Específicos

Objetivo Específico No. 1
Recolectar información de los últimos 2 años concerniente a los clientes <ul style="list-style-type: none">- Obtener datos de uso por aplicación usando como criterio de búsqueda el número de identificación del suscriptor.- Obtener datos comerciales y de facturación usando como criterio de búsqueda el número de identificación del suscriptor.- Obtener datos de servicio al cliente usando como criterio de búsqueda el número de identificación del suscriptor. (1 meses)

Objetivo Específico No. 2
Consolidar la información y organizarla. <ul style="list-style-type: none">- Determinar las 5 aplicaciones más usadas- Determinar las 5 herramientas más usadas- Determinar las 10 funcionalidades más usadas por aplicación y herramienta- Determinar los 10 contenidos más visitados en los últimos 6 meses- Determinar los 20 clientes que más usan las aplicaciones- Discriminar los 3 que más usan las aplicaciones clientes por sector económico- Discriminar los 3 clientes que más usan las aplicaciones por ubicación geográfica- Discriminar los 20 clientes más importantes por número de suscripciones- Discriminar los 50 clientes más antiguos- Determinar las 10 aplicaciones y/o herramientas con mayor número de quejas (1 meses)

Objetivo Específico No. 3

Analizar la información consolidada y organizada.

- Determinar los productos de los clientes que redujeron o cancelaron su suscripción
- Determinar los productos de los clientes que aumentaron o adquirieron suscripción
- Determinar los 5 paquetes de productos que más se venden o solicitan los clientes discriminándolos por sector económico
- Determinar el porcentaje en que se vio afectada la suscripción históricamente

(2 meses)

Objetivo Específico No. 4

Realizar una investigación de mercado en la cual

- Se realice una encuesta a 100 suscriptores del sector más representativo en ventas.
- Se aplique una encuesta a los 10 suscriptores más representativos por sector.
- Se aplique una encuesta a los 5 suscriptores con mayor y menor uso.

(1 meses)

10.2. Metodología

El procedimiento a seguir es el siguiente:

- Definir los reportes deseados
 - o Determinar los campos y requerimientos de información que conforman los diferentes reportes e informes
 - o Determinar los elementos gráficos que complementen los datos reportados
 - o Realizar varias iteraciones para la generación del documento de requerimientos de reportes con las diferentes áreas y unidades de negocio relacionadas
- Establecer la infraestructura sobre la cual se va a montar esta propuesta.
 - o Validar con infraestructura software, hardware disponible para el montaje de la solución
 - o De ser necesario cotizar elementos faltantes o nuevos requeridos
- Recolectar información
 - o Validar completitud de la información
 - o Validar estructura de la información
 - o Garantizar datos llave para cruce con otras estructuras de datos

- Garantizar que cumplan con los requerimientos de información iniciales
- Extraer la información recolectada y consolidarla en un único repositorio
 - Extraer información de servicio al cliente
 - Extraer la información de ERP de la organización
 - Financiera
 - Segmentación
 - Extraer la información comercial
 - Extraer la información de usos de aplicación
 - Google Analytics
 - Particulares de aplicación
 - Scripting de captura
 - Esta información debe ser historia a 3 años
- Normalizar la información obtenida (Transformación)
- Definir las relaciones de los datos con tablas de dimensión y definir los campos calculados
- Definición de cubos cubriendo los requerimientos de información inicial
- Proponer los métodos de análisis iniciales sobre la información final ya consolidada, normalizada y relacionada.

10.3. Presupuesto general del proyecto

Presupuesto aproximado en miles de pesos				
Rubro	UPB	Financiación propia	Otra institución	Total
Personal	17800	0	0	17800
Equipos	145	0	0	145
Materiales	100	50	0	150
Software	0	1	0	1
Bibliografía	600	493	0	1093
Viajes	0	0	0	0
Total	18645	544	0	19189

En cuanto software se usará la versión de SQL 2016 Developer la cual en el ambiente intranet es de costo cero (0) debido al licenciamiento corporativo de la organización.

11. Viabilidad financiera

Para DATASOL los clientes se pueden clasificar según el número y tamaño de sus suscripciones en:

- Grandes. Suscripciones anuales mayores a COP \$30.000.001
- Medianos. Suscripciones anuales entre COP \$5.000.001 y \$30.000.000
- Pequeños. Suscripciones anuales menores a COP \$5.000.000.

Se parte del hecho de que el nivel de deserción corporativo es del 2.3% y con la probabilidad de que los clientes pueden ser de cualquier tamaño se puede afirmar que el 0.8% pertenece a cada tipo de suscriptor. Si se toma el un valor promedio mensual puede haber un nivel de deserción o reducción en su suscripción que equivale aproximadamente a dejar de recibir COP\$ 45.000.000 por cada 100 clientes.

Minimizar estas cifras hace que el proyecto sea viable ante una inversión inicial de aproximadamente COP\$ 20.000.000 en un escenario optimista durante 6 meses en comparación al costo que tiene perder uno o varios clientes representados en el porcentaje de deserción o reducción mensual.

En un escenario menos ideal el costo se elevaría a COP\$ 25.000.000 siendo esta una cifra aún muy por debajo del efecto que tiene perder clientes.

Ya en un escenario muy pesimista el costo se elevaría a COP\$ 30.000.000 que al igual que el escenario anterior sigue siendo esta una cifra muy por debajo del costo generado por la reducción de suscripciones o deserción de clientes.

12. Plan de actividades

#	Tipo de actividad	Actividad	Mes 1				Mes 2				
			Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Semana 6	Semana 7	Semana 8	
1	Investigación teórica	Aprender teoría de conjuntos									
2		Conocer algunos algoritmos de comparación en SQL Server									
3		Aprender los conceptos básicos para la definición e implementación de cubos - Dimensiones - Medidas									

1 7	Definir reportes por paquetes de productos																		
1 8	Definir reporte de deserción histórico																		
1 9	Definir reporte de suscripciones nuevas vs deserción																		
2 0	Definir reporte de ingreso al mercado de productos tanto nuevos como tradicionales																		
2 1	Realizar encuestas a cliente de sectores representativos.																		
2 2	Consolidar encuestas a cliente de sectores representativos y proponer mejoras, cambios y nuevos productos y servicios																		

13. Plan de adquisiciones

- Adquirir lo identificado en el rubro de Bibliografía lo cual incluye tres (3) capacitaciones relacionadas a temas como cubos, BI y minería de datos en SQL server.
- Adquirir información de apoyo y complemento a las capacitaciones recibidas y/o adquiridas
- Adquirir un servidor para instalar SQL Server para el repositorio de los archivos origen de datos y posterior procesamiento de los datos
- Adquirir espacio de almacenamiento digital para almacenar el histórico de cargues y la información procesada
- Adquirir la suscripción y elaborar las entrevistas a los diferentes nichos de clientes
- Enviar la convocatoria a los diferentes usuarios para la resolución de las encuestas

14. Plan de riesgos

Ventajas

- La unificación de los datos por parte de áreas interdisciplinarias al interior de la compañía.
- Una visión más global de como son los clientes de DATASOL.
- La rectificación de la información ya que en ocasiones se vendía algo y se facturaba internamente el mismo producto, pero con otro tipo de suscripción y menos cantidad de usuarios.

- Identificar que un cliente que es bueno según la perspectiva de un área puede ser un mal cliente para otra área. Al cruzar información se puede evidenciar el nivel de cliente que representa para la organización.
- Contar con un número adecuado de informes de distribución pública al interior de la compañía con información transversal a todas las áreas que la conforman.
- Al ser un desarrollo interno e incluir a grupos interdisciplinarios según el detalle en cuestión, hacen que el conocimiento del proyecto se quede al interior de la compañía y no en una sola persona.
- Alinea a la compañía en pro de su misión y visión y genera una retroalimentación para reforzar estos elementos o fortalecerlos.
- La inversión extra es mínima ya que se hace con recursos propios de la compañía con una dedicación parcial, permitiendo el avance y gestión de otras actividades por fuera de este proyecto.

Desventajas

- Que las diferentes áreas no proporcionen toda la información requerida en el momento adecuado
- Que las diferentes fuentes de datos sean muy heterogéneas sin ningún criterio para su unificación y consolidación
- El nivel de certeza de los algoritmos de comparación puede generar actividades extras al cronograma para su correcto ejecutar.
- Que las personas que realicen la capacitación técnica se retiren de la compañía o sean asignados a otro cargo.
- Que los usuarios no resuelvan las encuestas propuestas por la compañía.

15. Plan de interesados

- Servicio al cliente requiere de información para realizar temas de implementación, capacitación, soporte y encuestas de satisfacción a los clientes, dando prioridad a aquellos cuya tendencia es la deserción, la disminución de suscripciones, la disminución de usos de aplicación (uso cero) y la cercanía a la fecha de caducidad de la suscripción. (Prioridad Alta)

- Comercial requiere segmentar a sus clientes para realizar mejores ofertas comercial acorde al fin comercial de cada cliente identificando costos por encima o por debajo de lo que debería ser la propuesta ideal. (Prioridad Alta)
- Financiera requiere garantizar que lo que se vendió sea lo mismo que se facturó tanto en cantidad como en tipo de venta. (Prioridad Media)
- Tecnología debe estar siempre alerta y dimensionar adecuadamente la infraestructura para atender el número de usuarios que han sido contratados de forma concurrente en las aplicaciones. Además, debe garantizar que la infraestructura interna a la compañía permita la generación y almacenamiento de los diferentes datos para los interesados y nombrados. (Prioridad Baja)
- Redacción contará con un insumo para invertir esfuerzos en los contenidos y herramientas de mayor interés para los clientes. (Prioridad Media)

16. Conclusiones y recomendaciones

Este proyecto solo es el inicio de una serie de análisis que se pueden realizar con respecto a la adquisición y uso de las aplicaciones y herramientas DATASOL.

Se evidencia que con una inversión comparativamente baja con respecto a lo que representa la deserción o disminución de las suscripciones por parte de los clientes se pueden madurar los productos, mejorar las estrategias comerciales y adecuar los métodos de retención de clientes, fortaleciendo los equipos de servicio al cliente y dándole herramientas al área comercial para ofertar de forma más certera los productos DATASOL.

Una vez la propuesta sea aceptada y se realice el proyecto en sí, se recomienda que la información tenga un seguimiento mensual como mínimo, debido a las reglas del negocio con respecto a las fechas de vencimiento de las suscripciones.

Incursionar en la minería de datos para extraer conocimiento útil, previamente desconocido, partiendo de grandes cantidades de datos desde distintos formatos y orígenes, permite que la organización evolucione y crezca como tal, no cada área independientemente sino evidenciando la sinergia que hay entre ellas mismas y que todas se alineen por un bien común.

17. Referencias

- Study on the visual experience of senior people shopping website interface
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85049218034&origin=resultslist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&relpos=1&citeCnt=0&searchTerm=#references>
- The performance analysis of distributed storage systems used in scalable web systems
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85053513915&origin=resultslist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&relpos=3&citeCnt=0&searchTerm=>
- Population prediction based on the multi-models and comparison research
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85053272589&origin=resultslist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&relpos=5&citeCnt=0&searchTerm=>
- Scenario-based insider threat detection from cyber activities
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85052700825&origin=resultslist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&relpos=17&citeCnt=0&searchTerm=>
- A model for inclusive design of digital libraries
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?origin=recordpage&zone=relatedDocuments&eid=2-s2.0-36348968360&citeCnt=0&noHighlight=false&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&relpos=0>
- LabBook: Metadata-driven social collaborative data analysis
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-84963754756&origin=relist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=>
- The dimension of age and gender as user model demographic factors for automatic personalization in e-commerce sites
<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85042207646&origin=resultslist&sort=plf->

[f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=bacc53e918c08912731179bd242437b9&sot=b&sdt=b&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=29&citeCnt=0&searchTerm=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85025821745&origin=resultslist&sort=plf-f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=bacc53e918c08912731179bd242437b9&sot=b&sdt=b&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=29&citeCnt=0&searchTerm=)

- Toward the development of a conventional time series-based web error forecasting framework

<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85025821745&origin=resultslist&sort=plf->

[f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=b926cfe1b2632aa1efb595a767ea1289&sot=b&sdt=cl&luster=scosubjabbr%2c%22COMP%22%2ct%2c%22SOC%22%2ct%2c%22ENGI%22%2ct&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=65&citeCnt=1&searchTerm=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85042929470&origin=resultslist&sort=plf-f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=b926cfe1b2632aa1efb595a767ea1289&sot=b&sdt=cl&luster=scosubjabbr%2c%22COMP%22%2ct%2c%22SOC%22%2ct%2c%22ENGI%22%2ct&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=65&citeCnt=1&searchTerm=)

- Characterizing user interest in NoSQL databases of social question and answer data

<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85042929470&origin=resultslist&sort=plf->

[f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=b926cfe1b2632aa1efb595a767ea1289&sot=b&sdt=cl&luster=scosubjabbr%2c%22COMP%22%2ct%2c%22SOC%22%2ct%2c%22ENGI%22%2ct&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=72&citeCnt=0&searchTerm=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85045065398&origin=resultslist&sort=plf-f&src=s&st1=website+statistics&nlo=&nlr=&nls=&sid=b926cfe1b2632aa1efb595a767ea1289&sot=b&sdt=cl&luster=scosubjabbr%2c%22COMP%22%2ct%2c%22SOC%22%2ct%2c%22ENGI%22%2ct&sl=33&s=TITLE-ABS-KEY%28website+statistics%29&relpos=72&citeCnt=0&searchTerm=)

- E-commerce big data computing platform system based on distributed computing logistics information

<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-85045065398&origin=resultslist&sort=plf->

[f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-84879098421&origin=reflist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=)

- The relative importance of usability and functionality factors for E-health web sites

<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-84879098421&origin=reflist&sort=plf->

[f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-22544443339&origin=reflist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=)

- What is usability in the context of the digital library and how can it be measured?

<https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-22544443339&origin=reflist&sort=plf->

[f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=](https://www-scopus-com.loginbiblio.poligran.edu.co/record/display.uri?eid=2-s2.0-22544443339&origin=reflist&sort=plf-f&src=s&st1=website++use+statistics&st2=&sid=8de75a41b0da60cfbd382e4257eec594&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28website++use+statistics%29&recordRank=)