

TABLA DE CONTENIDIO

BRIEF DEL RETO.....	1
a) Reto.....	1
b) Objetivo Estratégico	1
c) Antecedentes	1
d) Descripción del problema	2
e) Publico objetivo	3
f) Impacto esperado	3
g) Restricciones.....	4
h) PDS.....	4

BRIEF DEL RETO

a) Retos

¿Cómo podríamos digitalizar, automatizar y hacer analítica a la información documental de los procesos de Ecopetrol para maximizar su valor como un activo estratégico de la compañía?

b) Objetivo Estratégico

(¿Cómo se alinea con la estrategia empresarial?)

Optimización de costos y disciplina de capital.
Modelo flexible y adaptable incrementando la eficiencia operacional.
Fortalecer la gestión de medio ambiente.

c) Antecedentes

(¿Qué ha sucedido que se está generando un problema? ¿Qué se ha realizado previamente en la compañía, hay algún proyecto en curso? Detalla las iniciativas ya realizadas que dan información de éxitos y fracasos).

La pandemia del COVID19 ha hecho notar la importancia de la gestión digital y la producción de la información digital. En el marco de la operación diaria de Ecopetrol cada

uno de los procesos de la empresa produce altas cantidades de información que no se está digitalizando.

La cantidad de archivos físicos va en aumento y cada vez se encuentran nuevos repositorios de información lo cual causa que las personas que van a acceder a la información no tengan la información disponible de inmediato.

Actualmente ECOPETROL tiene estandarizada la gestión de la información de forma física, lo cual le ha permitido iniciar la etapa de transición de lo físico a lo digital. Se ha identificado información física para digitalización: 200 millones de imágenes lo que equivale a 200-250 mil cajas de información.

ECOPETROL está trabajando en una iniciativa llamada “Información y documento electrónico de archivo” (IDEA) que tiene los siguientes pilares:

- Correspondencia y sede electrónica: Flujo de firma – DocuSign.
- Correspondencia física: Transición de cartero a correspondencia digital y correspondencia física con distribución electrónica y física.
- Automatización de inventarios.

d) Descripción del problema

(Explicar con detalles la situación, para que sea muy evidente que el problema realmente es un problema y que hay una gran oportunidad si se soluciona. Datos, cifras, porcentajes, que le dan relevancia para invertir en la situación)

Se ha identificado la necesidad de transformar el servicio de gestión documental a un servicio de gestión de la información debido a que no se está aprovechando la información de los documentos que se generan en los procesos de ECOPETROL.

La información que se genera de los procesos tienen diferentes tipos: video, audio, texto, email e imágenes.

De igual manera, se ha identificado la existencia de varios repositorios y aun así hay información de procesos que no se encuentra en ninguno de estos y solo se encuentran en los documentos en físico.

Para lograr la gestión de la información se hace clave un proceso que nos permita de principio a fin lo siguiente:

- Producción de información digital.
- Digitalización inteligente de la información.

- Datificación: Procesamiento y normalización de la información.
- Automatización de flujos de información y tareas de gestión.
- Analítica de contenido para identificar patrones, lecciones y extracción de contenido y síntesis de la información.

e) Publico objetivo

(¿Quién es el destinatario de la solución?)

Funcionarios de Ecopetrol S.A
Gerencia Servicios Compartidos.
Entes de control internos (VCU, auditoria interna, control disciplinarios)
Entes de control externos (Fiscalía, Contraloría)

f) Impacto esperado

(¿Qué resultados espero obtener?, datos cuantitativos y cualitativos que ayuden a entender lo que se espera para el público objetivo y para la empresa)

Se espera que la solución digital permita digitalizar, automatizar y hacer analítica de la información de todos los procesos de ECOPETROL.

La solución debe facilitar:

- Producción digital de información y generación de valor a través de automatización de flujos y analítica de contenido.
- La toma de decisiones basado en análisis de datos.
- Monetizar la información de los documentos para generar valor real.

La solución tendrá los siguientes impactos:

- Convertir la información en un activo estratégico de la empresa.
- Eliminación de la intermediación humana en la gestión de la información.
- Aumentará la productividad.
- Reducción de costos y tiempos de ciclos de procesos asociados a la producción y búsqueda de la información.
- Reducción de riesgos de fuga, perdida de información y multas por falta de oportunidad y calidad en la información.
- Identificación de patrones y tendencias de la gestión para generar mejoras y eficiencia operativa.
- Reducción de costos de preservación y procesamiento de la información.
- Reducción de crecimiento archivo físico.
- Autogestión para acceder a la información.

- Evidencia de los registros de los procesos de ECOPETROL.
- Apoyará a mejorar la inteligencia y valor colectivo de la empresa.
- Mejorar la experiencia de usuario.

g) Restricciones

(Barreras que se puedan identificar para el desarrollo e implementación de la solución).

h) PDS

(Compendio de requerimientos o características que debe tener la solución)

Repositorio centralizado

Implementar sobre Opentext Extended ECM (xECM) flujos documentales que garanticen una gestión documental centralizada y unificada que permita clasificar, asociar metadatos, indexar, OCR para búsqueda sobre contenido, búsquedas amigables por nombre, contenido y metadatos, búsquedas avanzadas, versionar y controlar las versiones, controlar el acceso, llevar historial de cambios y accesos de los documentos, con el fin de facilitar auditorías y soportar la toma de decisiones de todos los procesos de Ecopetrol y su grupo empresarial, de manera uniforme, eficiente, segura e inteligente.

Asegurar la organización jerárquica del sistema de administración documental y taxonomía corporativa, alcanzando una correcta categorización de manera ascendente o descendente de la información, logrando búsquedas óptimas a través de la indexación de los metadatos propios del documento, ayudando al usuario a identificar los registros y la información de manera más eficiente.

Espacios digitales – content estándar para todas las áreas de Ecopetrol. (620 espacios según estructura orgánica funcional), con los estándares de gestión documental necesarios para que sea aplicado a todas las áreas de Ecopetrol y su GEE que se vayan incorporando a la solución documental.

Automatización de la gestión de producción digital de información

Automatización de la gestión, tanto flujos de producción digital de documentos como los de catalogación y gestión documental.

La clasificación automática debe hacer uso de técnicas de Inteligencia Artificial para ordenar y asignar los documentos, así mismo la digitalización inteligente incorpora el documento electrónico a flujos automatizados que eliminan la gestión manual en el proceso intervenido.

La automatización de la gestión se enfoca en la incorporación de modelos automáticos que soportan el ciclo de vida del documento en el proceso.

Digitalización.

Disponer de mecanismos de digitalización inteligente, digitalización certificada y digitalización confiable, que permita capturar, almacenar e indexar datos y documentos para disponerlos en flujos automatizados que habiliten la eliminación de la gestión manual del documento. El proceso de captura inteligente incluye indexación del documento, reconocimiento del documento y extracción de datos.

La digitalización certificada es un proceso que otorgará al documento electrónico resultante (copia electrónica certificada más sus correspondientes metadatos y firma asociada al proceso de digitalización) el mismo valor que el documento original en papel, lo que permitirá poder eliminar este soporte. Se establecerá a través de una instrucción si esta documentación en papel necesita un tiempo de almacenamiento previo a su eliminación, así como la destrucción de esta, ajustándose a la normativa interna de Ecopetrol para tal fin.

Buscador inteligente.

El Sistema de búsqueda que sea capaz de conectarse e indexar los repositorios de información y que se constituya como único punto de acceso a la información (documentos, audios, videos etc.) El cual soporte:

- Indexación que agilice la búsqueda de información y ayude a seleccionar con precisión la información requerida por los usuarios.
- Estándares de conexión con nuevas fuentes de información.
- Búsquedas basadas en contenido.
- Búsquedas basadas en conocimiento.

Analítica de contenido

Se requiere que el componente de analítica soporte los siguientes requerimientos funcionales:

- o Integración con el buscador
- o Soportar analítica de texto como función que permita ser invocada de forma parametrizada y desacoplada desde diferentes puntos.

- Análisis de lenguaje natural en texto líquido y tradicional para interpretación de contenido para responder preguntas del propósito, significado, y Datificación de los mensajes de datos.
- Análisis de sentimiento sobre los contenidos de información.
- Reconocimiento de contenido y su significado
- Extracción de contenidos específicos.
- Identificación del significado de contenidos.
- Entendimiento del lenguaje y extracción avanzada de contenido, entendimiento de la semántica del mensaje de datos por contexto, clase, tipología y área de negocio.
- Interpretación de imágenes y su contenido (OCR /ICR / OMR)
- Determinación de criticidad de los mensajes de datos.
- Verifique la Integridad de contenido de los documentos.
- Validación de contenidos correctos e información faltante.

Lineamientos de arquitectura

- El repositorio documental oficial es Opentext, por tanto todos los documentos reposan allí y si se requiere llevarlos a la nube para procesamiento masivo se realiza a un área de staging de un datalake.
- Todos los metadatos extraídos por los flujos analíticos y de machine learning deben quedar almacenados como metadatos asociados al documento en opentext o en un datamart con el modelo semántico que relacione estos datos con otro tipo de información.
- Para modelos desarrollados y administrados por ECOPETROL las herramientas son las de Microsoft. El espacio de trabajo para procesamiento de datos y desarrollo con R, Python, Scala Spark es Azure Databricks. Los modelos de ML se publican en Azure ML y se despliegan en Azure kubernetes services.
- Se puede incorporar la analítica y machine learning como servicio SaaS conectado a Opentext.