

**OPTIMIZACIÓN DE GESTIÓN DE PROYECTOS DE DESARROLLO COMUNITARIO CON GOBIERNO DE DATOS**

**Abstract**

El desarrollo rural sostenible es fundamental para mejorar la calidad de vida en comunidades marginadas, sin embargo, la falta de un gobierno de datos eficaz obstaculiza este proceso. Este proyecto propone la implementación de un gobierno de datos para una organización dedicada a proyectos comunitarios rurales. La recopilación de datos obtenida manualmente y de fuentes dispersas, conlleva a inconsistencias, falta de integridad y dificultad en el análisis. Un gobierno de datos permitirá estructurar normativas y responsabilidades claras para la gestión, recopilación, almacenamiento y uso de los datos, lo cual incluye recabar de forma eficiente y precisa la información para planificar y ejecutar los proyectos.

Dentro del proyecto, se propone estandarizar procesos y herramientas para la captura de datos, simplificar la carga administrativa, facilitando la comparación y el análisis de los mismos. Además, el establecer políticas de seguridad y privacidad de datos, protegerá la información sensible de manera efectiva, mitigando riesgos o mal uso de los datos.

La implementación de este proyecto también apoya la oportunidad para aumentar la credibilidad y la visibilidad de la organización ante potenciales colaboradores, con información completa, transparente, auditable gestionada con responsabilidad, permitiendo financiar de forma confiable sobre los proyectos brindados a los participantes de cada comunidad.

**Objetivo**

El objetivo de este proyecto está basado en mejorar la eficiencia operativa y el impacto social de la organización optimizando la gestión de proyectos de desarrollo comunitario.

Implementar un gobierno de datos efectivo, para garantizar calidad, seguridad y privacidad de los datos utilizados.

Establecer una cultura de datos sólida que promueva la toma de decisiones basada en datos y la transparencia en la gestión de proyectos.

Agregar gobierno de datos a la solución propuesta es crucial para garantizar la calidad, integridad, seguridad y privacidad de los datos, lo cual incluye la construcción de:

* Políticas y Normativas: que rijan el manejo de los datos en la organización, incluyendo la definición de roles y responsabilidades, procesos de gestión de datos y directrices para la privacidad y seguridad de la información.
* Catálogo de Datos: que describa de manera detallada los conjuntos de datos utilizados en la gestión de proyectos, incluyendo el origen, significado, calidad y uso permitido.
* Calidad de los Datos: implementar medidas para garantizar la calidad de los datos, como la validación automática en los formularios, limpieza periódica de los datos y estandarización de formatos.
* Seguridad y Privacidad: establecer controles robustos para proteger datos sensibles, controles de acceso basado en roles y monitoreo de actividades dentro la gestión.
* Cumplimiento Normativo: asegurar el cumplimiento de las normativas planteadas, respetando la protección de los datos, una manejo adecuado y seguridad de la información.
* Capacitación y Concientización: capacitar al personal técnico que realizará la captura de los datos por formularios, de los proyectos de portafolio, las políticas y prácticas a emplear de acuerdo con el gobierno de datos establecido, promoviendo una cultura de responsabilidad y cumplimiento.
* Auditoría y Monitoreo: establecer procesos de supervisión del cumplimiento de las políticas del gobierno de datos, identificando posibles riesgos y tomar medidas correctivas donde fuere necesario.

**Antecedentes de la Fundación**

La organización se dedica a apoyar proyectos de desarrollo comunitario en diversas comunidades, con el objetivo de promover la autosuficiencia de las familias beneficiarias. La gestión de proyectos actual se basa en procesos manuales y la falta de un gobierno de datos efectivo dificulta la optimización de los recursos y la medición del impacto de los proyectos.

El proceso existente incluye:

* Recopilación manual de datos en formatos de Excel.
* Falta de estándares de acuerdo con la calidad de los datos.
* Ausencia de una visión completa y actualizada del progreso de los proyectos por participante y comunidad.
* Riesgo de seguridad y privacidad de los datos.
* Falta de análisis determinado del total de suministros entregados y producción de generada por participante (Total de aves entregadas, número de muertes y nacimientos, producción actual y costos asociados).

Esto conlleva a perder la visibilidad operativa de acuerdo con el inventario, por semana y mes, respecto de cada comunidad activa.

Los procesos iniciales han optado por la integración de muchas fuentes, cada una de ellas obtenidas por cada técnico a cargo.

**Situación actual de la problemática**

El proceso actual incluye la visita a los participantes de cada proyecto por comunidad, cada técnico toma datos en papel y luego los traslada en hojas de cálculo al supervisor para integrar por proyecto y participante el seguimiento obtenido por comunidad. La ejecución de este trabajo se consolida en forma mensual para inspeccionar el avance de los proyectos. Si los datos contienen error, se consulta nuevamente al técnico a cargo y este posterior al seguimiento puede aplicar nuevos valores a los parámetros registrados.

Cada comunidad se registra activamente de acuerdo con un representante, líder y responsable de recibir al técnico asignado en su visita. Cada participante está asociado a un proyecto, por comunidad donde actúa como responsable de la producción, cuidado y gestión de los insumos recibidos. Dentro de este proceso se dificulta el control por área, al no contar con un estándar en los formatos de captura, para el registro de cada seguimiento. Por otro lado, se carece de la calidad en que los datos puedan ser reales, al permitir actualizaciones en un tiempo posterior a su a la visita del participante.

En cuanto a la producción y avance de proyectos, se dificulta el seguimiento total del número de muertes y nacimientos, respecto de los animales, no llevando un control específico por etapas y desconociendo parámetros importantes como las causas y criterios del ambiente.

No todos los proyectos cumplen con un formato estándar, por lo que el análisis de sus variables, dependen de diferentes fuentes, ajustadas a un único formato que se consolida manualmente para determinar su avance.

**Descripción de la cultura de datos actual**

La cultura de datos actual en la organización es limitada, con un enforque reactivo

en la gestión de datos y una falta de conciencia sobre la importancia del gobierno de datos y

la toma de decisiones basada en datos.

Se necesita concientizar al colaborador de la responsabilidad del tratamiento de sus datos en cada proceso, desde la captura operativa de los seguimientos hasta la consolidación y análisis de estos.

Aplicar política de cambios en los procesos de captura y diseño de las variables que la planificación requiera de acuerdo con los objetivos esperados, normativas de seguridad que no permitan desviar la calidad de la información obtenida y mantener la transparencia de los datos desde la supervisión en campo.

Valida los formatos elaborados de los formularios previos al seguimiento y visita con cada uno de los participantes, los productos a configurar bajo un catálogo estructurado.

Evidenciar la toma de los registros en fecha correspondiente, sin alteración o cambios en un proceso posterior.

Identificar los catálogos de información compuestos por los datos maestros y metadatos, roles administrativos definidos para garantizar la correcta operación.

**Situación esperada para la cultura de datos**

Se espera que el personal involucrado en este proyecto participe de acuerdo con los principios de gobernanza de datos. Esto incluye la integridad y transparencia. Practicar la comunicación efectiva con el personal involucrado del proceso, ser veraces en la información presentada. Ser claros con las instrucciones dadas, por parte de los responsables del proceso y claros en los informes obtenidos de seguimiento.

Se espera establecer una cultura de datos sólida que promueva la colaboración en el manejo de los datos, fomentando la toma de decisiones basada en datos, por cada uno de los niveles.

Que los supervisores de cada proyecto puedan obtener datos auditables, acompañados con documentación electrónica, como soporte para los requerimientos basados en auditoría,

Respaldar las decisiones en datos, procesos y controles establecidos, que apoyen a cada uno de los involucrados en la gestión.

Consolidar una fuente estandarizada de información que nos permita introducir controles, analizar los balances del proyecto y la tecnología empleada, analizar fenómenos obtenidos y estudio acerca de los datos recolectados.

Poder establecer quien tomara las decisiones, con que datos y cuando respecto de que situaciones y bajo que método de apoyo. Quien tiene la autorización de construir el seguimiento, el personal que planifica y coordina las visitas por proyecto, respecto de su orquestación.

**Estrategia de Datos Por Seguir**

A continuación, se presenta la estrategia de datos propuesta para el proyecto, los pasos son los siguientes:

* Se realizarán reuniones con los principales interesados para identificar y priorizar las problemáticas que el análisis de datos podrá resolver.
* Se definirá claramente qué información será necesaria para abordar estas problemáticas.
* Se determinarán las fuentes internas y externas de datos que se utilizarán, incluyendo los archivos de Excel existentes en el proceso actual.
* Se establecerán los requisitos técnicos y de negocio para la integración de estas fuentes de datos.
* Se desarrollará una plataforma web y móvil para administración y la captación de los datos del proceso en campo.
* Se construirán Data Marts específicos para casos de uso concretos.
* Se diseñará la arquitectura de un Data Warehouse utilizando un esquema de tipo estrella para facilitar las consultas analíticas.
* Se estimará el volumen de datos a manejar y se preverá la escalabilidad del almacenamiento en la nube (Oracle Cloud).
* Se crearán Dashboards interactivos según las necesidades de los usuarios finales.
* Para la visualización y análisis de datos, se utilizará PowerBI con conexión a Oracle Cloud y Pentaho.
* Para la limpieza y transformación de datos (ETL) se implementarán scripts de bases de datos en diferentes procesos administrados por Pentaho.
* Se implementarán mejores prácticas de gobernanza de datos para administrar los permisos y accesos.
* Se planificarán sesiones de capacitación para usuarios y administradores del sistema en el uso de las herramientas y la interpretación de los datos.
* Se establecerá una línea temporal para estas capacitaciones, asegurando que coincidan con los lanzamientos de fases del proyecto.

**Arquitectura de la Información Planteada**

La arquitectura del proyecto propuesto, abarca los componentes y las interacciones entre ellos, considerando las herramientas y tecnologías mencionadas.

Como primer paso en esta propuesta de arquitectura se desarrollará una plataforma web en Ruby On Rails para la captación y administración de datos a través de interfaces web, la plataforma estará configurada para interactuar directamente con una base de datos de tipo relacional preparada para la Fundación. El siguiente paso es el desarrollo de una aplicación móvil en Android para permitir la entrada de datos en tiempo real desde dispositivos móviles, la app móvil utilizará APIs para enviar datos al servidor central. Como siguiente paso se utilizará la herramienta Pentaho Data Integration para la creación de los siguientes procedimientos:

* JobOrigenes para conectar y extraer datos proporcionados por la plataforma Ruby on Rails y la aplicación Android.
* JobStage y JobDW para manejar estos nuevos formatos de datos entrantes y asegurar su correcta transformación y carga.

Otro paso importante es la automatización a través de Cron en el servidor Linux de la ejecución del ETL para la carga de los datos a la estructura de los Datamarts en el esquema en estrella para adaptarse a los nuevos tipos de datos y volumen proveniente de las aplicaciones en los horarios de 6:00 am y 11:00 pm todos los días con la opción de ejecución manual a demanda.

Como paso final se realizará la construcción e integración de dashboards interactivos con las estructuras de datos para maximizar el rendimiento de las consultas basadas en los datos más recientes de la plataforma web y móvil.

**Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente**

**Matriz RACI**

Para la definición de la matriz RACI, se han identificado los siguientes roles:

* Alta Gerencia: Contempla a los directivos de la Fundación, quienes tienen el poder de decisión del proyecto
* Equipo de proyecto: Personal experto de Fundación separado estratégicamente para acompañar en las diferentes etapas del proyecto.
* TI: Equipo de Tecnología e Información a cargo del desarrollo e implementación de la propuesta.
* Equipo Operativo: Técnicos, supervisores, y participantes claves de los diferentes proyectos comunitarios, que apoyarán en el traslado de información y pruebas de las plataformas a desarrollar.

Basado en los roles, se desarrolla la siguiente matriz RACI, identificando las actividades y las responsabilidades de cada rol.



**Productos Digitales**

Debido a que en Fundación el proceso de captura de datos es manual realizado en papel y digitado en Excel, esta propuesta genera como productos digitales los siguientes:

Una plataforma web completamente funcional diseñada para la entrada y gestión de datos, que proporciona interfaces para el ingreso manual de datos, así como para la supervisión y administración de estos, dentro de sus beneficios se tiene la centralización de datos con acceso controlado, la mejora en la eficiencia de la captura de datos al eliminar el uso de Excel para ingresos manuales y la disponibilidad de informes en tiempo real y seguimiento de la entrada de datos.

Una aplicación móvil diseñada para facilitar la entrada de datos desde cualquier ubicación, optimizando los procesos que requieren movilidad o acceso remoto, dentro de sus beneficios se tiene la flexibilidad y acceso a la plataforma desde cualquier lugar, aumentando la cobertura de datos, captura de datos en tiempo real, lo que reduce los retrasos en el procesamiento y análisis de datos, interfaz fácil de usar que puede aumentar la participación del usuario y la precisión de los datos.

Un conjunto de procesos ETL (extracción, transformación y carga) adaptados para integrar datos desde la plataforma web y la aplicación móvil hacia el Data Warehouse en la nube, dentro de sus beneficios están la automatización del procesamiento de datos que asegura consistencia y reduce errores humanos, capacidad para manejar grandes volúmenes de datos y complejidades en la transformación y mejoras en la velocidad y la eficiencia de los procesos de actualización de datos.

Un repositorio centralizado en la nube que almacena datos estructurados en Data Marts siguiendo un esquema en estrella, optimizado para análisis y reportes, con los siguientes beneficios:

* Escalabilidad y seguridad mejoradas proporcionadas por soluciones en la nube.
* Capacidad de realizar análisis complejos y generación de reportes con alta disponibilidad.
* Integración eficiente con herramientas de BI para análisis avanzados.

Un conjunto de Dashboards y reportes configurados en PowerBI para visualizar y analizar datos desde el Data Warehouse con visualización avanzada de datos para una mejor toma de decisiones, interactividad y capacidad de personalización que permite a los usuarios profundizar en los datos específicos y acceso remoto a información crítica de negocio, mejorando la agilidad empresarial de Fundación.

Los productos digitales estarán regidos de acuerdo con las normas y procedimientos establecidos en el gobierno de datos, garantizando la integridad, seguridad y privacidad de los mismos.

Estos productos digitales no solo modernizan la infraestructura existente sino también agregan valor al negocio al mejorar la accesibilidad, la calidad de los datos y la eficiencia de los procesos empresariales de Fundación. Además, abren nuevas oportunidades para análisis más profundos y respuestas más rápidas a las tendencias del mercado

**Tecnología por utilizar para la explotación de los datos**

Basado en la arquitectura propuesta se utilizarán las siguientes tecnologías para la explotación de datos:

- Plataforma Web: Ruby On Rails.

- Aplicación Móvil: Android.

- Extracción, Transformación y Carga: Pentaho Data Integration.

- Base de datos relacional: PostgreSQL en infraestructura Oracle Cloud.

- Repositorio Data Warehouse: Infraestructura Oracle Cloud.

- Servidor: Linux Ubuntu Server, infraestructura Oracle Cloud.

- Presentación de Datos: PowerBI Profesional.

**Evaluación Económica**

A continuación, se presenta una tabla con las diferentes observaciones en cada costo asociado.

|  |  |  |  |
| --- | --- | --- | --- |
| **Actividad** | **Costo Unitario** | **Costo Proyecto (12 meses)** | **Costo / Observación** |
| Costo por infraestructura de servidor (hardware y software) – Plataforma Oracle Cloud | $190 x mes  7.80 tasa cambio  Q1,482 x mes | $2,280 x proyecto  7.80 tasa cambio  Q17,784 x proyecto | Asumido por la Organización al tratarse de espacio en la nube que ya está siendo pagado como parte del presupuesto normal. |
| Costo por software a utilizar para la infraestructura de servidor, y equipo de desarrollo. | $0 | $0 | Se utilizará software gratuito, por lo cual no hay ningún costo asociado.  El equipo para utilizar en el desarrollo de la aplicación es otorgado por la Organización donde se llevará a cabo esta solución. |
| Costo por licenciamiento de software de visualización | PowerBI Profesional Anual $127  Tasa 7.80  Q990.60 Anual | 1 Licencia para desarrollador  1 Licencia para consulta  Q1,981.20 | La empresa ya tiene estas licencias, por lo cual no se incurre en ningún gasto adicional. |
| Costo por equipo de móvil y planes de datos | Q1000 por equipo  60 usuarios aproximados  Q60,000 x equipos  Q150 plan de datos al mes  Q9000 x plan de datos mensual | Q60,000 x equipo  Q108,000 x plan de datos x proyecto | Se utilizará el equipo móvil empresarial, no es necesario el incremento en cantidad de equipo o pago adicional por datos consumidos.  La Organización ya está asumiendo los costos por proyectos propios de movilidad por lo cual no hay costos asociados. |
| Costo asociado por el desarrollo de la aplicación | Q0 | Q0 | El costo por el desarrollo de la aplicación es asumido al ser personal de Fundación quien desarrollará la aplicación. |
| **Total, Proyecto (+)** |  | **Q187,765.20** |  |
| **Costos asumidos por la Organización (-)** |  | **Q187,765.20** |  |
| **Costo de inversión** | | **Q0** |  |

Se concluye en la evaluación económica que no se incurre en ningún gasto debido a que todos los costos son asumidos por la Organización donde se llevará a cabo la implementación de la solución, y el costo asociado al desarrollo de la aplicación es asumido por el personal de TI de la Organización como parte interesada de llevar a cabo esta investigación.

**Conclusiones**

* La implementación de un Gobierno de datos apoyará a la eficiencia operativa de la Fundación, contribuyendo significativamente a garantizar la calidad, seguridad y privacidad de los datos, creando una gestión más eficaz y transparente de los proyectos gestionados por la organización, impactando de manera positiva a los proyectos sociales al poder medir y ajustar las estrategias en tiempo real.
* Mejora en la eficiencia para la toma de decisiones informadas, transparencia y confiabilidad en los procesos, lo que permitirá que las partes interesadas comprendan y apoyen a las diferentes iniciativas de la Fundación.
* Efectividad en un gobierno de datos mediante la introducción de políticas y normativas, estándares de calidad de datos, medidas de seguridad y privacidad, estableciendo un marco que respalda la integridad y la gestión de los datos.
* Concientización a todo el personal involucrado, sea técnico o administrativo, sobre la importancia del manejo adecuado de los datos de manera responsable y conforme a las políticas establecidas, asegurando la integridad y confidencialidad de estos.
* Establecimiento de un monitoreo y auditoria continua de la gestión de datos que apoya a la identificación rápida de cualquier desviación o riesgo potencial, contribuyendo a mantener altos estándares de calidad y seguridad en las actividades de la organización.

**Recomendaciones**

* Mantener un ciclo de capacitación continua para el personal involucrado en la gestión de datos, no solo a los técnicos de los proyectos, sino también a los gestores y líderes, enfocándose en mantener y mejorar la gestión de datos.
* La herramienta de análisis de datos debe mantener una constante evolución, para ello es importante que se capacite a personal que conoce el proceso de la Fundación con el objetivo de tener a los actores correctos para el desarrollo de las visualizaciones necesarias que apoyen a la toma de decisiones basada en datos.
* Implementar un proceso de gestión del cambio para los diferentes niveles de personal con que cuenta la Fundación, derivado a los cambios en la manera en que se obtendrán los datos, cambios tecnológicos, y en la manera en que se analizará la información, haciendo énfasis en la importancia de mantener y fortalecer el gobierno de datos por implementar.
* La alta dirección debe demostrar un compromiso continuo, desde la creación de las actividades relacionadas al gobierno de datos, hasta la implementación y cumplimiento de estos, ofreciendo apoyo y guiando en cada etapa de la implementación de este proyecto.