

ANEXO TÉCNICO NO. 8 – SERVICIO ESPECIALIZADO DE ANALITICA DE DATOS.

El presente anexo describe los procedimientos, características técnicas, exigencias y requisitos para la provisión de los servicios requeridos contemplando los siguientes aspectos:

- **Antecedentes**
- **Servicios y requerimientos para analítica de datos**
- **Análisis de requerimientos y diagnóstico para analítica de datos**
- **Implementación y construcción del lago de datos (data lake) o bodegas de datos y servicios de interoperabilidad**
- **Extracción, limpieza, transformación y cargue de información de los sistemas al lago de datos (data lake) o bodegas de datos**
- **Resultados de la analítica de datos**
- **Metodología de desarrollo**
- **Acuerdos de nivel de servicio**

1. ANTECEDENTES

ENTerritorio en el marco de sus proyectos tecnológicos y de arquitectura empresarial ha priorizado la implementación de los sistemas de información asociados a la operación y gestión sus líneas de negocio así:

- El sistema ERP (por sus siglas en inglés, enterprise resource planning) para la gestión y administración de los módulos de:
 - o Contabilidad y Pagaduría
 - o Compras y Facturación
 - o Costos
 - o Contratación
- El sistema de información CRM por sus siglas en inglés (por sus siglas en inglés, customer relationship management) para la gestión de sus clientes en los módulos de:
 - o Análisis de Mercado
 - o Gestión Comercial
 - o Direccionamiento Comercial
 - o Satisfacción del Cliente
- El sistema Hub de Proyectos.
- El sistema de vinculación de clientes.
- El sistema de gestión documental Orfeo.
- Fuentes de datos de Sistemas misionales, financieras, operativos, entre otros que estén en procesos de producción o en desarrollo y se requieran en el modelo de analítica.

Los sistemas de información ERP y CRM se desarrollaron bajo la arquitectura de Software como Servicio – SaaS y en la solución Microsoft Dynamics de Nube Publica Azure manejando información estructurada desde sus portales Web.

Los sistemas FVC, Orfeo y HUB de Proyectos se encuentran bajo arquitecturas OnPremise.

A continuación, se relaciona la información de cada sistema de información y la conformación global de la base de datos que lo soporta:

Sistema de Información	Año de Puesta en Producción	BD y Tamaño promedio de BD en GB	Tipo de Información	Numero de Tablas o Vistas
ERP	2022	No aplica: Modelo SaaS.	Estructurada	13.728 tablas 6.769 vistas
CRM	2021	No aplica: Modelo SaaS.	Estructurada	10 tablas
HUB de Proyectos	2020	Motor Azure Sql Server Tamaño 0.3 Gb	Estructurada	104 tablas
FVC	2018	Motor Oracle 12. Tamaño 0.2 Gb	Estructurada	45 tablas
Orfeo	2016	Motor Oracle 12. Tamaño 17 Gb	Estructurada	100 tablas

Bajo estos elementos descritos se solicita que el contratista implemente los requerimientos de analítica de datos descritas en los siguientes numerales:

2. ANALISIS DE REQUERIMIENTOS Y DIAGNOSTICO PARA ANALITICA DE DATOS

ENTerritorio requiere que el contratista realice un diagnóstico técnico y diseño de arquitectura para la implementación de los modelos analíticos de datos sobre la información de sus sistemas de información descritos que se requieran en el desarrollo del modelo, que permiten el desarrollo de su misión.

Así mismo, se determina la necesidad que el contratista entregue a los usuarios de la entidad herramientas con las que ofrezca la analítica de datos, para sacar conclusiones, identificar alertas, efectuar control en tiempo real, hacer enfoques descriptivos y predictivos de la gestión de su información y poder aprovechar al máximo la información guardada en repositorios de sus sistemas de información.

Para el diseño del modelo de analítica, el contratista debe trabajar con la información que se genere de manera histórica en los sistemas de información acorde a la línea de tiempo.

Las actividades que debe ejecutar el contratista en este ítem son:

1. Hacer el diagnóstico de capacidades tecnológicas para la implementación de analítica de datos y requisitos de sostenibilidad para el periodo de tiempo en el plazo de ejecución del contrato.



2. Recolectar la información, aplicar entrevistas y encuestas, u otro mecanismo que requiera al personal que determine el Contratista y/o ENTerritorio, para establecer clases de analítica de datos a aplicar, esfuerzo asociado a los casos de uso documentados y seleccionar los casos de uso a implementar.
3. Documentar los requerimientos incluyendo costos de operación por servidores, almacenamiento, licenciamiento, seguridad, herramientas estadísticas y de predicción; migración, presentación de información, capacitación, soporte, mantenimiento y asesoría.
4. Diseñar del modelo de analítica de datos para ENTerritorio.
5. Desarrollar e implementar de los casos de uso de analítica de datos para ENTerritorio que se requieran en el marco del análisis realizado.
6. Diseñar tableros de control, flexibles y portables (responsive) con base en la operación y resultados del modelo de analítica de datos.
7. Ejecutar la transferencia de conocimiento a los usuarios que determine ENTerritorio, que no podrá ser inferior a 30 personas.

El contratista debe cumplir con los entregables del diagnóstico y requerimientos de analítica de datos así:

1. Documento técnico específico y detallado con el diagnóstico de la analítica de datos para ENTerritorio, el modelo de analítica y los casos de uso requeridos y priorizados.
2. Documento de Diseño de reportes dinámicos, flexibles y portables (responsive) con base en la operación y resultados del modelo de analítica de datos.
3. Documento de evidencia de la transferencia de conocimiento a los usuarios que determine ENTerritorio.

Los documentos deben ser entregados a la supervisión del contrato para su aprobación, a los 3 días hábiles de cumplirse la ejecución o entrega de actividad a satisfacción.

3. IMPLEMENTACION Y CONSTRUCCIÓN DEL LAGO DE DATOS (DATA LAKE) o BODEGAS DE DATOS Y SERVICIOS DE INTEROPERABILIDAD

De acuerdo con el modelo diseñado de Analítica de Datos, se requiere diseñar, configurar e implementar uno o varios lagos de datos (Data Lake) o Bodegas de datos u otro medio de almacenamiento masivo que permita almacenar datos estructurados, semiestructurados y no estructurados, proveniente de los sistemas transaccionales, Servicios Web, APIs, hojas electrónicas, archivos PDF, entre otros.

Las actividades que debe ejecutar el contratista en este ítem son:

1. Diseñar, configurar e implementar el Lago de Datos (Data Lake) o bodegas de datos que permite el cargue y la actualización de diferentes tipos de datos estructurados, semiestructurados, no estructurados, de las diferentes fuentes de datos: sistemas transaccionales, hojas electrónicas, archivos .PDF, entre otros. Para la transformación y tratamiento de la información no estructurada se podrán utilizar herramientas de “machine learning” que estén soportadas por el fabricante.
2. Crear todos los objetos de base de datos que conforman el Lago de Datos o bodegas de datos como tablas, vistas, paquetes, funciones, disparadores,



secuencias, esquemas y demás que sean necesarios para el desarrollo y soporte de sus funcionalidades.

3. Realizar todos los ajustes inherentes para la instalación de la base de datos.
4. Desarrollar componentes informáticos para la actualización de la información del lago de datos (Data Lake) o bodegas de datos, desde los sistemas transacciones y demás fuentes de información, según la frecuencia que determine la entidad.
5. Desarrollar las APIs y servicios web (WS) que se requieran para garantizar la interoperabilidad entre los sistemas transaccionales, el lago de datos (Data Lake) o bodegas de datos, fuentes de datos no estructurados y fuentes externas.
6. Aplicar mecanismos para el respaldo de datos y seguridad del Data Lake o bodegas de datos.
4. Ejecutar la transferencia de conocimiento a los usuarios que determine ENTerritorio, que no podrá ser inferior a 5 personas.
7. Documentar de forma precisa y detallada los procesos implementados de interoperabilidad y el diseño e implementación del Lago de Datos (Data Lake) o bodegas de datos conforme a los estándares internacionales en la materia y los lineamientos del MinTIC.
8. Documentar la base de datos tanto a nivel de instalación y configuración como a nivel de diseño. Para ello deberá documentar y elaborar el Modelo Entidad Relación, Catálogo de datos y demás documentación relacionada.

Los DataLake o bodegas de datos debe implementarse y operar en la modalidad de Nube Publica Azure bajo el tenant de ENTerritorio.

Las licencias requeridas y sus costos asociados de soporte para la operación de la solución integrada de analítica de datos estarán a cargo del contratista. La solución integral puede estar soportada por herramientas Open Source.

El contratista debe cumplir con los entregables así:

1. Documento técnico de arquitectura de los Data Lake o bodegas de datos, Catálogo o diccionario de datos, modelo Entidad Relación, artefactos de software y elementos tecnológicos que conforman su almacenamiento y demás documentación relacionada a la base de datos.
2. Documento técnico de arquitectura APIs, Servicios Web y demás artefactos de software y elementos tecnológicos que conforman la solución de interoperabilidad con los Data Lake o bodegas de datos.
3. Licenciamiento y soporte aplicado para la operación de los Data Lake o bodegas de datos, por el término de plazo del contrato.
4. Documento de evidencia de la transferencia de conocimiento a los usuarios que determine ENTerritorio.

Los documentos deben ser entregados a la supervisión del contrato para su aprobación, a los 3 días hábiles de cumplirse la ejecución o entrega de actividad a satisfacción.

4. EXTRACCIÓN, LIMPIEZA, TRANSFORMACIÓN Y CARGUE DE INFORMACIÓN DE LOS SISTEMAS AL LAGO DE DATOS O BODEGAS DE DATOS



El contratista deberá desarrollar, configurar y poner en funcionamiento las conexiones bidireccionales de las fuentes origen de los datos que alimentarán los modelos desarrollados mediante un proceso de extracción, limpieza, transformación y cargue de datos (ETL).

El contratista deberá garantizar la aplicación automática y permanente de algoritmos para calidad del dato como fase previa al procesamiento por parte de los reportes e informes que respondan a cada necesidad.

Los resultados de la aplicación de calidad de dato deben permitir generar reportes de estado de la data utilizada con el fin de servir como insumo de mejora para los procedimientos o sistemas que originan los datos entrantes.

Los reportes de calidad del dato deben tener como mínimo: nombre de la fuente de información, variable afectada, error detectado y fecha de detección. Para ello el contratista desarrollará, configurará y pondrá a punto las tareas y procedimientos de limpieza y calidad del dato.

El proceso de extracción, integración y preparación deberá ser automático, sin necesidad de intervenciones mecánicas, humanas o manuales.

Los procesos ETL de datos no deben utilizar consultas (SQL) directas a las bases de datos.

El contratista deberá garantizar que no se sobrepase el 5% de la carga transaccional común, para lo cual, deberá realizar un testeo de estrés documentado y entregado al supervisor del contrato, que permita determinar la mejor configuración recomendada; en este sentido, la comunicación entre la solución y las fuentes no debe generar pérdida de funcionalidad ni sobrecarga transaccional sobre el sistema origen que afecte su funcionamiento. El supervisor validará el estado óptimo para aprobación.

Las actividades que debe ejecutar el contratista en este ítem son:

1. Extraer los datos de los sistemas de información estructurados y no estructurados que gestiona ENTerritorio descritos en este documento.
2. Cargar el resultado de extracción, limpieza y transformación para alimentar el o los Data Lake o bodegas de datos que se requieran mediante herramientas ETL o equivalentes.
3. Aplicar la limpieza, estandarización y controles en las fuentes de datos que garanticen su calidad. Debe estandarizar la información y establecer controles para garantizar datos confiables desde la fuente.
4. Ejecutar la transferencia de conocimiento a los usuarios que determine ENTerritorio, que no podrá ser inferior a 5 personas.
5. Entregar la documentación precisa y detallada sobre los procesos implementados de interoperabilidad y el diseño e implementación del Lago de Datos (Data Lake) o bodegas de datos conforme a los estándares internacionales en la materia y los lineamientos del MinTIC.

6. Realizar acompañamiento a los administradores técnicos de ENTerritorio o Externos para las configuraciones que se requiera en las bases de datos origen.

La solución integral para los procesos ETL debe operar en ambiente de Nube Publica Azure bajo el tenant de ENTerritorio.

Las licencias requeridas y sus costos asociados de soporte para la operación de la solución integrada de analítica de datos estarán a cargo del contratista. La solución integral puede estar soportada por herramientas Open Source.

El contratista debe cumplir con los entregables así:

1. Data Lake o bodegas de datos con información cargada tras el proceso de ETL.
2. Desarrollo de componentes informáticos ETL para la actualización de la información del lago de datos (Data Lake) o bodegas de datos.
3. Desarrollo de las APIs que se requieran para asegurar la interoperabilidad entre los diferentes Sistemas de Información determinados para analítica de datos.
4. Desarrollo de componentes API y Servicios Web (interfaces automáticas) utilizados en la construcción del Data Lake.
5. Documento técnico preciso y detallado de arquitectura del Data Lake o bodegas de datos y de las interfaces automáticas.
6. Ejecutar la transferencia de conocimiento a los usuarios que determine ENTerritorio, que no podrá ser inferior a 5 personas.

Los documentos deben ser entregados a la supervisión del contrato para su aprobación, a los 3 días hábiles de cumplirse la ejecución o entrega de actividad a satisfacción.

5. RESULTADOS DE LA ANALÍTICA DE DATOS

El contratista deberá diseñar, implementar, presentar y configurar los resultados de la analítica de datos mediante una interfaz de acceso Web en la herramienta de BI (PowerBI); sin embargo, se podrá proponer a la supervisión del contrato otra opción, quien una vez revisadas las condiciones de favorabilidad funcional y sostenibilidad podrá dar aval para incorporarlo. Esta posible opción no tendrá costo adicional para ENTerritorio.

El contratista deberá utilizar para el desarrollo de resultados de la analítica de datos la entrega de reportes e informes y la utilización de lenguajes de programación “open source (Python y R)”, sin embargo, en caso de requerirse algún complemento que no cumpla esta característica y sea pertinente incorporar, deberá presentarse la(s) opción(es) al supervisor del contrato, quien una vez revisadas las condiciones de favorabilidad funcional y sostenibilidad podrá dar aval para incorporarlo. La incorporación de este componente no tendrá costo adicional para ENTerritorio.

La solución implementada por el contratista debe permitir la generación reportes dinámicos con interfaz gráfica interactiva para la gestión de filtros de la información reportada y con la opción de personalización y exportación en formatos de ofimática como Excel, PDF, Word o visualizados en línea.

De igual forma, la solución debe permitir creación de nuevos reportes personalizados por parte de los usuarios que sean capacitados en los lenguajes de programación y herramientas implementadas para su auto sostenimiento.

La interfaz gráfica debe ser amigable, de fácil usabilidad y accesibilidad y acordada con el supervisor del contrato

Todos los desarrollos por entregar deberán contar con una interface gráfica de acuerdo con el formato de comunicación e imagen de ENTerritorio.

Si el contratista usa para la construcción componentes de software o librerías propietarias que requieran licenciamiento, deberá entregar a ENTerritorio las licencias de aquellas que sean empleadas en la arquitectura de la solución, estas licencias deberán ser de uso ilimitado en tiempo, a perpetuidad, y entregadas por el contratista mediante informe y recibo a satisfacción por el supervisor del contrato.

El contratista deberá diseñar los modelos de presentación grafica de la solución que el usuario final visualizará guardando los lineamientos de identidad corporativa.

La aplicación deberá funcionar completamente en los principales navegadores web en las versiones más recientes.

Los reportes e informes que entregue el contratista deben contener las siguientes características:

1. Debe ser una interfaz gráfica que combine texto y elementos visuales haciendo énfasis en la parte gráfica, esta debe tener un alto contenido de elementos visuales que ayuden a transmitir un mensaje, entender una situación o significado más rápido y fácil, que con tan sólo usar texto.
2. Se debe poder ver toda la información del Cuadro de Mando (NO Scroll), para entender la información eficientemente y en todo momento.
3. Debe contener mecanismos visuales, concisos e intuitivos.
4. Debe ser personalizado, según las necesidades de la organización, área o el objetivo para el cual está diseñado.
5. Deben poder monitorear uno o más objetivos según sea el caso de uso.
6. Debe apoyar a la toma de decisiones.
7. Deben utilizar alertas, enfocadas principalmente en alarmar o informar sobre una situación particular y tomar acciones en el momento oportuno.
8. Debe viabilizar la comparación de tendencias y diversas variables así estas no estén relacionadas.
9. Deben permitir un análisis desde varios niveles de la entidad (gerencial, funcional, operativo, entre otros).
10. Deben permitir comparaciones a gran escala.
11. Debe ofrecer funciones como filtros, búsqueda de tablas o archivos, agregación, cálculos matemáticos y estadístico.
12. Debe permitir generar notificaciones por medios como el correo electrónico.

6. METODOLOGÍA DE DESARROLLO

El proyecto debe realizarse bajo un marco de trabajo para desarrollo ágil de software que permita evidenciar objetivos cortos y tangibles, así como retroalimentación constante para corrección temprana, de igual forma se deberá incluir metodología de analítica de datos como ejemplo CRISP-DM.

El contratista deberá entregar al inicio del contrato (máximo 15 días posterior a la firma del acta de inicio) un cronograma detallado de las actividades, coordinado y avalado con el supervisor del contrato.

Todo complemento y desarrollo incorporado en el proyecto, será propiedad de ENTerritorio.

El contratista deberá garantizar el desarrollo del proyecto global, las actividades de ejecución y cierre del requerimiento (entrega de la documentación final, formalización de actas de recibo a satisfacción, encuesta de satisfacción y certificación de experiencia)

Una vez la solución se encuentre en producción, el contratista deberá cargar todos los códigos fuentes, desarrollos, ajustes y configuraciones de toda la solución en el servidor de versiones de la entidad.

En caso de ser componentes el contratista deberá indicar de forma explícita marca y versión del componente adjuntando, suministrar el soporte y compatibilidad del fabricante, y discriminación mediante informe de forma clara y explícita del licenciamiento utilizado y suministrado a ENTerritorio.

7. ACUERDO DE NIVEL DE SERVICIO

NOMBRE ANS	DISPONIBILIDAD DEL SERVICIO DE ANALITICA DE DATOS
DEFINICIÓN:	<p>La indisponibilidad es el número total de minutos, durante el mes contratado, en los que el servicio o componente o desarrollo o implementación o configuración no se presentó a la Entidad, dividido en el número total de minutos en el mes contratado.</p> <p>La medición la hace el contratista monitoreando permanentemente el servidor durante el mes.</p> <p>Los resultados del monitoreo son mantenidos por el contratista para que puedan ser consultados por ENTerritorio en cualquier momento durante la duración del contrato. La información mantenida por el contratista le debe permitir a ENTerritorio verificar la disponibilidad histórica del servicio en los meses anteriores y durante el mes en curso.</p>
MEDICIÓN:	$\left(1 - \frac{\text{Número total de minutos en que el servicio no está disponible}}{\text{Número de días en el mes contratados} \times 24 \text{ horas} \times 60 \text{ minutos}}\right) \times 100$ <p style="text-align: center;">DISPONIBILIDAD EXIGIDA \geq 99.98 %</p>



PENALIDAD	PENALIDAD POR NO CONFORMIDAD		DESCUENTO
	$99.90\% \leq \text{Disponibilidad} < 99.98\%$		10%
	$99.80\% \leq \text{Disponibilidad} < 99.90\%$		20%
	$99.70\% \leq \text{Disponibilidad} < 99.80\%$		50%
	$\text{Disponibilidad} < 99.70\%$		100%
APLICA A:	Facturación de la línea de Analítica de Datos (Valor antes de IVA)		

NOMBRE ANS	OPORTUNIDAD DE ENTREGABLES DEL SERVICIO DE ANALITICA DE DATOS
DEFINICIÓN:	La oportunidad es el número total de entregables recibidos a satisfacción por la supervisión del contrato, durante el mes, dividido el número total de entregables acordados en el mes.
MEDICIÓN:	$\frac{\text{número total de entregables recibidos a satisfacción durante el mes}}{\text{número total de entregables acordados en el mes}}$
PENALIDAD	<i>0% a 49.9%, descuento del 10% del valor mensual facturado. 50% a 59.9%, descuento del 7% del valor mensual facturado. 60% a 69.9%, descuento del 5% del valor mensual facturado. 70% a 79.9%, descuento del 3% del valor mensual facturado.</i>
APLICA A:	Facturación de la línea de Analítica de Datos (Valor antes de IVA)

1. TRANSICION, CIERRE Y ENTREGA

Cuatro (4) meses antes de la finalización del contrato el Contratista debe entregar el plan de Transición de salida. Este plan se debe ejecutar durante los tres (3) últimos meses de contrato:

Esta etapa se desarrolla en paralelo con la etapa de operación y no exime al Contratista de los descuentos por incumplimiento de ANS y del desarrollo normal de la operación.

Las actividades y entregables asociados a la transición del servicio son los documentos referenciados como entregables en este anexo actualizados al corte del día de finalización del contrato.