

Requerimiento Técnico

ESTRATEGIA DE CALIDAD DE LA INICIATIVA DATOS ABIERTOS

1.- Descripción de la necesidad que la entidad pretende satisfacer

El Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC), de acuerdo con la Ley 1341 de 2009 modificada por la Ley 1978 de 2019, desarrolla políticas y planes enfocados a las tecnologías de la información y las comunicaciones que constituyen un componente vital para el crecimiento y desarrollo del sector, con el fin de brindar acceso a toda la población en el marco de la expansión y diversificación de las TIC.

De conformidad con lo dispuesto en el artículo 21 de la Ley 1978 de 2019, el objeto del Fondo Único de Tecnologías de la Información y las Comunicaciones es *“financiar los planes, programas y proyectos para facilitar prioritariamente el acceso universal y el servicio universal de todos los habitantes del territorio nacional a las Tecnologías de la Información y las Comunicaciones, garantizar el fortalecimiento de la televisión pública, la promoción de los contenidos multiplataforma de interés público y cultural, y la apropiación social y productiva de las TIC, así como apoyar las actividades del Ministerio de Tecnologías de la Información y las Comunicaciones y la Agencia Nacional del Espectro, y el mejoramiento de su capacidad administrativa, técnica y operativa para el cumplimiento de sus funciones”*.

Ahora bien, con el objeto de incrementar la efectividad del Ministerio de Tecnologías de la Información y las Comunicaciones (MINTIC) en lo que se refiere al diseño y formulación de políticas que coordinen y estandaricen acciones de implementación y uso de las Tecnologías de la Información y las Comunicaciones en el Gobierno Nacional, el Presidente de la República expidió el Decreto 1414 de 25 de agosto de 2017, a través del cual modificó la estructura del Ministerio de Tecnologías de la Información y las Comunicaciones y creó el Despacho del Viceministro de Economía Digital, al cual pertenece la Dirección de Gobierno Digital, y tiene entre otras, las siguientes funciones: *“1. Analizar y proponer directrices de tecnologías de la información que cumplan los parámetros requeridos en materia de información estatal, de seguridad y protección de la información, coordinando con las entidades pertinentes en los temas de su competencia. (...) 3. Planear y liderar la definición de estándares y estructura tecnológica necesaria para el manejo de entornos de información compartida que garanticen la interoperabilidad de los sistemas del Gobierno, asegurando la disponibilidad, calidad, confiabilidad y seguridad de la información por parte de las*

entidades que los administran. (negrita fuera de texto) (...) 9. Promover la cooperación entre las autoridades nacionales, regionales y locales, así como las relaciones con instituciones comunitarias, en materia de Gobierno en Línea. **10. Elaborar e implementar las políticas necesarias para promover y facilitar que el Estado ponga a disposición de la ciudadanía en general los conjuntos de datos que permitan desarrollar procesos de innovación y emprendimiento de base tecnológica mediante el uso de las Tecnologías de la Información y las Comunicaciones.** (negrita fuera de texto)”

Adicionalmente, el Decreto 1008 del 14 de junio de 2018, por el cual se establecen los lineamientos generales de la política de Gobierno Digital y se subroga el capítulo 1 del título 9 de la parte 2 del libro 2 del Decreto 1078 de 2015, Decreto Único Reglamentario del sector de Tecnologías de la Información y las Comunicaciones, establece que son propósitos de la Política de Gobierno Digital: “4.1. *Habilitar y mejorar la provisión de servicios digitales de confianza y calidad.* 4.2. *Lograr procesos internos, seguros y eficientes a través del fortalecimiento de las capacidades de gestión de tecnologías de información.* **4.3. Tomar decisiones basadas en datos a partir del aumento el uso y aprovechamiento de la información.** 4.4. *Empoderar a los ciudadanos a través de la consolidación de un Estado Abierto* 4.5. *Impulsar el desarrollo de territorios y ciudades inteligentes para la solución de retos y problemáticas sociales a través del aprovechamiento de las TIC”.*

En virtud del propósito de la Política de Gobierno Digital “**Tomar decisiones basadas en datos a partir del aumento en el uso y aprovechamiento de la información**” la Dirección de Gobierno Digital busca mejorar la toma de decisiones por parte de la entidad, ciudadanos, usuarios y grupos de interés, para impulsar el desarrollo de servicios, políticas, normas, planes, programas, proyectos o asuntos de interés público, a partir del uso y aprovechamiento de datos que incorporan estándares de calidad y seguridad en su ciclo de vida (generación, recolección, almacenamiento, procesamiento, compartición, entrega, intercambio y eliminación).

Este propósito tiene su marco legal base en la **Ley 1712 del 06 de marzo de 2014 “Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional”**, donde los Datos Abiertos cobraron un impulso mucho mayor dentro de la agenda de transparencia del país, dado que se estableció el principio de máxima publicidad según el cual toda información en posesión, bajo control o custodia de un sujeto obligado es pública y no podrá ser reservada o limitada sino por disposición constitucional o legal, de conformidad con la ley. El derecho de acceso a la información genera la obligación correlativa de divulgar proactivamente la información pública y responder de buena fe, de manera adecuada, veraz, oportuna y accesible a las solicitudes de acceso por parte de los ciudadanos, lo que a su vez conlleva la obligación de producir o capturar la información pública.

A través del DECRETO 103 del 20 de enero del 2015, se reglamenta parcialmente la Ley 1712 de 2014 y se dictan otras disposiciones, en el Artículo 11. Publicación de Datos Abiertos, se especifica lo siguiente: *“Las condiciones técnicas de que trata el literal k) del artículo 11 de la Ley 1712 de 2014 para la publicación de datos abiertos, serán elaborados por el Ministerio de Tecnologías de la Información y las Comunicaciones y publicados en el Portal de Datos Abiertos del Estado colombiano o la herramienta que lo sustituya.”*, de igual manera los Datos Abiertos, se definen como *“todos aquellos datos primarios o sin procesar, que se encuentran en formatos **estándar e interoperables que facilitan su acceso y reutilización**, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos”*.

Adicionalmente a través del CONPES 3920 de 2018 política nacional de explotación de datos se establecieron metas de uso y aprovechamiento de datos que de acuerdo a los diagnósticos en dicho conpes a pesar de que Colombia ocupa una posición muy destacada en los índices internacionales de datos abiertos en la región, presenta bajo puntaje de avance en implementación, **lo cual se ha visto afectado por la disponibilidad real de datos abiertos del Gobierno, así como la calidad de los datos que sí están disponibles.**

Es por ello que surge la necesidad de implementar estrategia de calidad con el fin promover el uso y explotación de los datos públicos del gobierno a partir del desarrollo de nuevos componentes de analítica y calidad, el acompañamiento técnico especializado a las entidades públicas, la generación de material y talleres de capacitación, para que se obtenga una mejora significativa de la calidad de los conjuntos de datos en el portal nacional datos.gov.co y de esta manera mejorar los servicios prestados a la ciudadanía por parte de las entidades, generar nuevos productos que den solución a las problemáticas públicas y se tomen mejores decisiones basadas en datos.

2.- Descripción del objeto a contratar, con sus especificaciones y la identificación del contrato a celebrar.

2.1. Descripción del Objeto a contratar

2.1.1 Objetivo general:

Implementar la estrategia de calidad de la iniciativa datos abiertos a partir de la identificación de oportunidades de mejora en la calidad de los datos abiertos del portal datos.gov.co, la creación de nuevos componentes de analítica y calidad, el acompañamiento técnico especializado y la generación de material y talleres de formación y capacitación.

2.1.2 Objetivos específicos:

1. Diagnosticar la calidad de los conjuntos de datos abiertos del portal datos.gov.co
2. Diseñar y desarrollar la estrategia para mejorar la calidad de los conjuntos de datos del portal datos.gov.co
3. Desarrollar nuevos componentes de analítica y calidad a la herramienta de calidad del portal datos.gov.co
4. Realizar acompañamiento técnico especializado a las entidades publicadoras en el portal datos.gov.co
5. Generar material y talleres de formación y capacitación de la iniciativa datos abiertos para las entidades públicas

2.1.3 Alcance:

A continuación, se describe el alcance del requerimiento:

Se debe hacer uso y mejora de la herramienta de calidad disponible en el portal para diagnosticar el estado de la calidad de los datos del portal <https://herramientas.datos.gov.co/es/calidaddedatos>

Se deben tener en cuenta los errores frecuentes de calidad ya identificados en el portal <https://herramientas.datos.gov.co/es/blog/conoce-el-nuevo-proceso-de-aprobación-antes-de-publicar-tus-datos>

Al igual que los lineamientos disponibles como:

Al igual que los lineamientos disponibles como:

- La guía nacional de datos abiertos <https://herramientas.datos.gov.co/es/blog/conoce-el-nuevo-proceso-de-aprobación-antes-de-publicar-tus-datos>
- La guía de calidad e interoperabilidad https://herramientas.datos.gov.co/sites/default/files/A_guia_de_estandares_final_0.pdf
- Inventario de datos del portal <https://www.datos.gov.co/dataset/Asset-Inventory-Public/uzcf-b9dh>
- Las cuales deben ser actualizadas de acuerdo con los lineamientos vigentes, a las recomendaciones que surjan de la implementación de la estrategia de calidad y a la imagen corporativa vigente
- Se debe automatizar los procesos de calidad actualmente disponibles en la iniciativa tales como las comunicaciones masivas
- Se debe realizar el acompañamiento a las entidades públicas con el fin de mejorar la calidad de los errores frecuentes identificados
- Se deben ejecutar estrategias de formación a las entidades con el fin de mejorar la calidad de conjuntos de datos priorizados en procesos tales como automatización de conjuntos de datos, consolidación de conjuntos, errores de metadata y de registros

2.1.4 Elementos o información a tener en cuenta:

La estrategia de calidad debe partir de los procesos previos utilizados en la iniciativa para la mejora de los conjuntos de datos del portal nacional de datos abiertos

- Herramientas de calidad <https://herramientas.datos.gov.co/es/calidaddedatos>
- Proceso actual de calidad <https://herramientas.datos.gov.co/es/blog/conoce-el-nuevo-proceso-de-aprobación-antes-de-publicar-tus-datos>
- Conjunto de datos de calidad <https://www.datos.gov.co/Ciencia-Tecnolog-a-e-Innovaci-n/Calidad-de-Datos-Abiertos/efv4-ibbr>
- Inventario de datos del portal <https://www.datos.gov.co/dataset/Asset-Inventory-Public/uzcf-b9dh>
- La guía nacional de datos abiertos <https://herramientas.datos.gov.co/es/blog/conoce-el-nuevo-proceso-de-aprobación-antes-de-publicar-tus-datos>
- La guía de calidad e interoperabilidad https://herramientas.datos.gov.co/sites/default/files/A_guia_de_estandares_final_0.pdf

- Lenguaje común de intercambio de información <http://lenguaje.mintic.gov.co/>

2.1.4 Entregables:

1. Diagnosticar la calidad de los conjuntos de datos abiertos del portal datos.gov.co

Analizar la base de inventarios de activos del portal nacional de datos abiertos haciendo uso de las herramientas de analítica disponibles para procesamiento de grandes volúmenes de información y apoyándose en la herramienta de calidad disponible en el portal y de visualización como Power BI.

- Inventario de activos del portal <https://www.datos.gov.co/dataset/Asset-Inventory-Public/uzcf-b9dh>
- Analizar e identificar oportunidades de mejora en el código de Diagnóstico de la herramienta de calidad de datos sobre los criterios de evaluación e implementar las mejoras necesarias sobre los problemas identificados. (valores negativos, cálculos imprecisos). Herramienta de calidad <https://herramientas.datos.gov.co/es/calidaddedatos>

Resultados esperados:

1. 1 Documento con diagnóstico de la herramienta de calidad actual y de la calidad de los conjuntos de datos donde se clasifiquen y prioricen los errores de calidad a abordar y la estrategia de acuerdo con los [lineamientos de calidad de la iniciativa](#)

1.2 Documento de Arquitectura de Datos Moderna para el procesamiento masivo de los conjuntos de datos del portal

1.3 Actualización de [conjunto de datos de errores frecuentes](#)

2. Desarrollar nuevos componentes en la herramienta de calidad del portal de datos abiertos

- Actualización de la herramienta de Diagnóstico de Calidad de Datos
 - Identificar mejoras en la herramienta de Diagnóstico de Calidad de Datos <https://herramientas.datos.gov.co/es/calidaddedatos>

- Implementar las mejoras en el código de la herramienta de calidad de datos sobre los criterios de evaluación y sobre los problemas identificados teniendo en cuenta los conceptos de almacenamiento masivo y grandes volúmenes de datos. (valores negativos, cálculos imprecisos) <https://herramientas.datos.gov.co/es/calidaddedatos>
- Proponer e implementar nuevas formas de presentar resultados visuales en la herramienta- Power BI tales como: histogramas, series, mapas de calor (Por ejemplo: <https://www.digitalmenta.com/wp-content/uploads/2018/07/Mapa-Calor-AdWords-Impresiones.png>) permitiendo clasificar por errores visualmente los conjuntos de datos.

<https://herramientas.datos.gov.co/es/calidaddedatos>
- Desarrollar un módulo adicional en [la herramienta de calidad de datos](#) para medir criterios de la calidad estadística introduciendo conceptos como:
 - Pertinencia: El grado en que las estadísticas satisfacen las necesidades de los usuarios actuales y potenciales.
 - Oportunidad y puntualidad: La velocidad con que se publican los datos en relación con lo que miden y el grado de adhesión de las actualizaciones de datos al calendario de publicación previsto. (Revisar librería LEILA [LEILA - Librería de calidad de datos | Datos Abiertos Colombia](#)) o (Phyton)

<https://herramientas.datos.gov.co/es/calidaddedatos>
- Diseñar e implementar una herramienta o motor de autodiagnóstico modelo Machine Learning para el posterior cargue de nuevos conjuntos de datos en el portal [datos.gov.co](#) de manera que la entidad productora del dato tenga detalle de la calidad de los datos y que dicha herramienta pueda proponer la normalización de datos y recomendaciones de calidad al respecto, la herramienta debe permitir enviar un informe de los resultados a el correo electrónico de la entidad.
 - Para esta herramienta se requiere el desarrollo del Front end (React/ Angular entre otras) embebida en el portal de Datos abiertos en el módulo de calidad

- Los scripts de Algoritmos y reglas podrían utilizar python, spark, pyspark, entre otros. Revisar librería [LEILA - Librería de calidad de datos | Datos Abiertos Colombia](https://herramientas.datos.gov.co/es/fichatecnicacalidad) <https://herramientas.datos.gov.co/es/fichatecnicacalidad>
- Tablero de monitoreo de monitoreo y seguimiento a la calidad en Power BI

Resultados esperados

- 2.1. Plan de pruebas y migración e informes de ejecución de pruebas y migración de la herramienta de calidad
- 2.2 Actualización del Documento de Arquitectura Moderna
- 2.3. Código Fuente con cesión de derechos
- 2.4. Módulo de Calidad Implementado y publicado en el portal de datos abiertos en el menú calidad
- 2.5. Manuales técnicos y de usuario final
- 2.6. Capacitación al equipo de datos abiertos y material de capacitación publicado en el portal (Videotutoriales, Presentaciones y Manuales de uso)
- 2.7. Carta de Garantía y soporte por 1 año

3. Actualizar las guías y material de apoyo de la iniciativa datos abiertos para las entidades públicas

Resultados esperados

- 3.1 Actualización de La guía de calidad e interoperabilidad
https://herramientas.datos.gov.co/sites/default/files/A_guia_de_estandares_final_0.pdf
- 3.2 Actualización de La guía de datos abiertos
<https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf>
- 3.2 Actualización y publicación de material de apoyo para la formación en analítica y calidad de datos (presentaciones, videos, piezas para redes, tutoriales y guías) sobre calidad, sello de excelencia y lenguaje común de intercambio
<http://lenguaje.mintic.gov.co/>

4. Diseñar y desarrollar la estrategia para mejorar la calidad de los conjuntos de datos

- Planear como se abordarán los diferentes tipos de errores identificados de manera masiva, los instrumentos, guías, videos que deben crearse y/o actualizarse, y cual sería estrategia de difusión y acompañamiento a las entidades a partir de talleres de formación, asesorías técnicas especializados y la estrategia para hacer monitoreo y medición de la mejora de la calidad
- Integrar la iniciativa Sello de excelencia
https://sellodeexcelencia.gov.co/documents/UTSF_SDE_Requisitos_de_calidad_para_datos_abiertos_2019_12_02_v_2_0.pdf

Resultados esperados:

- 4.1. Documento de plan de mejora de la calidad con las estrategias a abordar, cronograma de trabajo, recursos a utilizar.
- 4.2. Mínimo 4 talleres nacionales virtuales de calidad (Debe incluir capacitación en el uso de la herramienta de calidad desarrollada).
- 4.3. Mínimo 50 entidades nacionales **de hoja de ruta** y 50 entidades territoriales con acompañamiento técnico especializado para mejorar la calidad de sus datos en temas de automatización o en los errores frecuentes identificados a nivel de metadata y registros (Debe incluir capacitación en el uso de la herramienta de calidad desarrollada) Ver documento de hoja de ruta <https://herramientas.datos.gov.co/es/blog/se-publica-para-comentarios-borrador-del-documento-hoja-de-ruta-de-datos-estrat%C3%A9gicos> Conjunto de datos de hoja de ruta <https://www.datos.gov.co/Ciencia-Tecnolog-a-e-Innovaci-n/Conjuntos-de-Datos-Abiertos-Estrat-gicos-Nacional-/ufnp-7sta/data>
- 4.4. Mínimo 4500 conjuntos de datos mejorados en calidad a nivel de metadata y/o estandarización con base a los lineamientos y soluciones a errores de calidad establecidos <https://www.datos.gov.co/Ciencia-Tecnolog-a-e-Innovaci-n/Errores-Frecuentes-de-Calidad-Conjuntos-de-datos-a/xbc7-65j4>
- 4.5. Informe de Implementación de estrategia de calidad

Nota: Ver los criterios de aceptación del Anexo 6

2.2 Especificaciones del contrato

El proveedor deberá entregar un informe de actividades y avance de ejecución del proyecto con una periodicidad semanal, y deberá comprometerse a presentar el cronograma de actividades en un archivo compatible con Microsoft Project.

La generación de las actas de las reuniones de seguimiento serán responsabilidad del proveedor.