

**PROGRAMA DE APOYO A LA MODERNIZACIÓN DE LA
DIRECCIÓN DE IMPUESTOS Y ADUANAS NACIONALES -
DIAN**

CONTRATO DE PRÉSTAMO BID 5148/OC-CO

**RFI – SERVICIOS DE DESARROLLO REPOSITORIO DE DATOS
DATAR**

SOLICITUD DE INFORMACIÓN

JUNIO DE 2026

TABLA DE CONTENIDO

1. TÉRMINOS Y SIGLAS	3
2. LISTADO DE ANEXOS	4
3. CRONOGRAMA.....	4
4. OBJETIVO DE LA SOLICITUD DE INFORMACIÓN	4
5. CONTEXTO DEL SERVICIO ACTUAL DE DATAR	5
6. ALCANCE	6
5.1 Tipos de productos considerados para la prestación de servicios	7
5.1.1 Productos de Información Usables	7
5.1.2 Productos técnicos	8
5.2 Enfoque técnico a considerar para la prestación de servicios	9
7. SOLICITUD DE INFORMACIÓN	13
6.1 Información General del Interesado.....	13
6.2 Capacidades Técnicas y Experiencia	14
6.3 Características de las Soluciones Tecnológicas Propuestas	14
6.4 Equipo Base del Proveedor	15
6.5 Aspectos Financieros (Estimaciones Preliminares)	15

1. TÉRMINOS Y SIGLAS

TÉRMINO / SIGLA	DESCRIPCIÓN
DataR	Repositorio Único de Datos de la DIAN.
Producto de información usable	Activo de datos implementado en el DataR y disponible en el ambiente de producción de la DIAN, una vez ha surtido las etapas de desarrollo, pruebas y validación establecidas, desarrollado en atención al alcance y a los criterios definidos en cada solicitud, bajo las condiciones técnicas y funcionales correspondientes. Este integra los procesos de transformación y actualización de la información, incorpora controles de calidad de datos, dispone de mecanismos de seguimiento y monitoreo, cuenta con la documentación técnica correspondiente y se encuentra estabilizado, con condiciones habilitadas para su uso, así como validado para su utilización por parte de los usuarios de la entidad.
Producto técnico	Conjunto de componentes, estructuras, procesos, automatizaciones y demás activos técnicos implementados en el DataR para habilitar, soportar y garantizar la integración, transformación, calidad, gobierno, trazabilidad, interoperabilidad, despliegue y aprovechamiento de los datos de la entidad. Incluye, entre otros, modelos de datos, flujos de integración y transformación, reglas de calidad, metadatos, mecanismos de trazabilidad, estándares técnicos, procesos de despliegue y gestión de cambios, así como demás componentes reutilizables requeridos para la sostenibilidad y evolución del repositorio.
Capacidades técnicas del DataR	Habilitadores tecnológicos y prácticas técnicas implementadas que permiten al DataR gestionar el ciclo de vida de los datos, soportando su integración, almacenamiento, procesamiento, transformación, calidad, seguridad, trazabilidad, interoperabilidad y aprovechamiento. Estas capacidades contribuyen a la evolución, madurez, sostenibilidad y fortalecimiento continuo del repositorio de datos de la DIAN.
Data Lakehouse	Arquitectura de datos que integra capacidades de Data Lake y Data Warehouse para almacenamiento, procesamiento y analítica.
Arquitectura Medallón	Modelo de organización de datos por capas Bronze, Silver y Gold para garantizar calidad y trazabilidad.
Unity Catalog	Servicio de gobernanza y catálogo de datos utilizado en Databricks para control de acceso, linaje y administración de metadatos.
XML	Lenguaje de marcado utilizado para el intercambio, estructuración y almacenamiento de información, ampliamente empleado en documentos electrónicos, integración de sistemas y procesamiento de datos estructurados y semiestructurados.
CDC (Change Data Capture)	Estrategia para identificar y procesar cambios realizados en las fuentes de datos.

CI/CD	Integración y despliegue continuos para automatización de pipelines y promoción entre ambientes.
-------	--

2. LISTADO DE ANEXOS

#	NOMBRE ANEXO
1	Anexo 1 – Información Interesado RFI DATAR.xlsx
2	Anexo 2 – Capacidades y experiencia del Interesado RFI DATAR.xlsx
3	Anexo 3 – Características Solución RFI DATAR.xlsx
4	Anexo 4 – Información equipo de trabajo RFI DATAR.xlsx
5	Anexo 5 – Cuadro de Precios RFI DATAR.xlsx

3. CRONOGRAMA

ACTIVIDAD	FECHA
Publicación y lanzamiento del RFI	23 de junio de 2026
Plazo para realizar preguntas y/o aclaraciones por parte de los interesados	10 de agosto de 2026
Fecha y hora límite para envío de respuestas al RFI por parte de los interesados	21 de agosto de 2026

4. OBJETIVO DE LA SOLICITUD DE INFORMACIÓN

En el marco del Programa de Apoyo a la Modernización de la DIAN, se adquirió en el año 2022 un proveedor para la implementación de los servicios de datos, para aportar en el cumplimiento del objetivo de “(...) mejorar la eficiencia de la gestión tecnológica, los datos y la seguridad de la información para optimizar la toma de decisiones y proteger la información”.

Como parte de la ejecución de este contrato, el proveedor realizó la habilitación de la solución tecnológica del DataR como repositorio centralizado de datos de la DIAN. Como resultado la DIAN actualmente cuenta con una plataforma tecnológica operativa, soportada en una arquitectura de tipo Data Lakehouse implementada en la nube primaria de Microsoft Azure. Esta plataforma constituye el repositorio central para la consolidación, gestión y explotación de los datos institucionales y consumo de datos de la entidad. Esta arquitectura dispone de capacidades habilitadas para la ingesta, transformación y almacenamiento de datos provenientes de diferentes sistemas de información de la DIAN, los cuales integran el repositorio central de datos. La implementación requirió la generación de procesos de gestión de datos orientados a la captura, validación y estructuración de la información, facilitando su disponibilidad para análisis de datos, interoperabilidad entre áreas y uso operativo.

En línea con lo anterior, el propósito de esta Solicitud de Información (RFI) es recopilar datos relevantes del mercado, identificar opciones disponibles, conocer las capacidades de los proveedores y obtener estimaciones preliminares de costos que, le permitan a la entidad la estructuración de un eventual proceso de contratación. Se busca especialmente comprender la forma en que los proveedores pueden ofrecer modelos de prestación del servicio bajo demanda, orientados al desarrollo de productos de información y/o de datos, garantizando condiciones de calidad, seguridad, mejora continua y gestión eficiente.

Estos servicios deberán permitir la ejecución de procesos de integración, transformación, estructuración, validación, disponibilidad, aprovechamiento de los datos y maduración de la arquitectura, mediante la construcción de productos usables, en concordancia con las políticas institucionales de gobernanza, calidad, seguridad de datos e información definidas por la entidad.

Para lograr este objetivo, la DIAN pone a disposición de los interesados el presente documento de Solicitud de Información (RFI – Request for Information), con el fin de recopilar insumos que faciliten el análisis frente a las alternativas de prestación del servicio y los enfoques disponibles en el mercado.

Por lo anterior, se invita a los interesados en participar en esta solicitud de información que servirá de base para un eventual proceso de contratación a responder de manera completa y detallada las preguntas de la **SECCIÓN 6. SOLICITUD DE INFORMACIÓN**, entendiéndose que su participación es fundamental para lograr un entendimiento claro de la prestación de los informados en este documento y sus anexos. Las respuestas recibidas no tendrán ningún tipo de relación o vínculo con el proceso de contratación de este servicio y la participación en este RFI no genera derecho, obligación ni compromiso para un eventual proceso de contratación.

5. CONTEXTO DEL SERVICIO ACTUAL DE DATAR

La DIAN cuenta actualmente con un Repositorio Único de Datos denominado DataR, el cual consiste en una plataforma tecnológica implementada sobre una arquitectura tipo Data Lakehouse con arquitectura medallón en la nube primaria de Microsoft Azure, que dispone de capacidades habilitadas para la ingesta, transformación, almacenamiento y disponibilidad de datos institucionales provenientes de diferentes sistemas de información de la entidad.

Esta plataforma constituye la arquitectura y repositorio central para la consolidación, gestión y aprovechamiento de la información, soportando procesos y productos para el consumo de datos por parte de las áreas misionales de la entidad. Ahora bien, es necesario mencionar que, teniendo en cuenta el marco de referencia del DAMA, el DataR actualmente se puede clasificar en un estado de madurez entre los niveles 2 (Gestionado) y 3 (Definido), por lo cual se requiere fortalecer las capacidades técnicas orientadas a la integración progresiva de nuevas fuentes de información, la consolidación de los procesos de calidad de datos, la estructuración de datos en niveles de consumo, y la construcción de productos de información que respondan a las necesidades

funcionales de la DIAN y a los cambios normativos y de negocio que puedan impactar la gestión de los datos.

Considerando las capacidades que deben fortalecerse, así como la evolución y madurez que se requiere alcanzar, se espera que las propuestas presentadas contemplen la capacidad de soportar cargas operativas de procesamiento de datos de alta exigencia, incluyendo escenarios de almacenamiento y procesamiento a escala de terabytes y petabytes, así como escenarios de procesamiento referencia de hasta 50 millones de registros diarios, con niveles adecuados de desempeño, escalabilidad, disponibilidad y seguridad del servicio. Esta referencia corresponde a un parámetro de dimensionamiento técnico para el diseño e implementación de productos y servicios de datos dentro del DataR, y representa escenarios potenciales de procesamiento que podrían presentarse durante la prestación del servicio.

En este contexto, los interesados deberán partir de las capacidades actualmente disponibles en el DataR, incluyendo la arquitectura Lakehouse, la plataforma de datos basada en Databricks y las capacidades de gobierno y gestión de datos habilitadas mediante Unity Catalog, las cuales actualmente soportan la organización y administración de datos a través de catálogos y esquemas, la gestión de accesos basada en grupos y roles, así como capacidades disponibles para trazabilidad y linaje de datos que podrán ser objeto de fortalecimiento y ampliación como parte de la evolución del DataR. Estas capacidades deberán ser consideradas como parte de la solución propuesta, definiendo, cuando aplique, las acciones requeridas para su fortalecimiento, consolidación o evolución tecnológica. Asimismo, se busca identificar propuestas que evidencien un valor agregado y un análisis de costo-beneficio favorable para la entidad en el mediano y largo plazo, contribuyendo a acelerar la entrega de productos de información de alto impacto.

6. ALCANCE

Con el propósito de integrar, gestionar y disponer los datos institucionales conforme a las necesidades de la DIAN, los servicios requeridos se enmarcan en la implementación y evolución de:

- a. Productos de información usables: activos de datos listos para consumo funcional, tableros, salidas, consultas, archivos, datasets o servicios de información, entre otros.
- b. Productos técnicos: pipelines, modelos, reglas de calidad, automatizaciones, metadatos, linaje, CI/CD, CDC, trazabilidad y componentes reutilizables, entre otros.

Estos se implementarán mediante estimaciones de esfuerzo que permitan definir la complejidad y tiempo de duración de cada desarrollo; asimismo estos servicios comprenden las actividades de migración, integración, transformación, estructuración y aprovechamiento de la información institucional, que serán ejecutados de manera planificada, priorizada y sostenible, y que resultan

necesarios para la consolidación, sostenibilidad, evolución, madurez y crecimiento del DataR, con un horizonte de ejecución proyectado hasta el año 2028.

5.1 Tipos de productos considerados para la prestación de servicios

5.1.1 Productos de Información Usables

En este contexto, el alcance contempla escenarios de carga operativa de procesamiento de datos de alta exigencia, considerando variaciones en la demanda, concurrencia de procesos y requerimientos de disponibilidad, desempeño, escalabilidad y seguridad, lo cual se enmarca en la implementación, evolución y mantenimiento de los productos de información que se construyan por parte del oferente, específicamente en lo relacionado con:

1. Conjuntos de datos integrados por dominio funcional (tributario, aduanero, facturación electrónica, exógena, entre otros), estructurados a partir de la consolidación de múltiples fuentes (bases de datos, esquemas, tablas y/o archivos), que contengan la información organizada de entidades como contribuyentes, declaraciones, documentos electrónicos y registros administrativos, disponibles para su consulta y explotación.
2. Soluciones parametrizables, configurables, reutilizables, auditables y escalables para la generación de información por dominio funcional, que permitan la consulta de datos a partir de criterios definidos por el usuario (como identificadores, periodos, variables, formatos o roles), soportados sobre estructuras y herramientas integradas en DataR.
3. Componentes para la generación de salidas de información para grandes volúmenes de datos, que permitan la obtención de datos en diferentes niveles de detalle o agregación, con posibilidad de exportación en formatos como archivos planos, conjuntos de datos estructurados para consumo y tableros/reportes de visualización.
4. Estructuras y servicios de datos de documentos electrónicos, construidos a partir del procesamiento de información estructurada y no estructurada (XML, repositorios de datos y servicios asociados), que permiten la consulta y generación de salidas de información.
5. Mecanismos de generación y/o recepción de información para convenios interinstitucionales, incluyendo cruces, validaciones y disposición de datos en formatos requeridos, con otras entidades del estado.

Se debe tener en cuenta que los productos de información y/o datos deben ser expuestos como activos reusables para diferentes dominios de negocio, teniendo en cuenta las especificaciones y dinámicas propias de cada dominio.

5.1.2 Productos técnicos

De manera complementaria a los productos de información orientados al consumo, se requiere la implementación y/o evolución de productos técnicos que permitan fortalecer las capacidades estructurales del DataR, facilitando la integración de nuevas fuentes, la estandarización de procesos y la sostenibilidad de la plataforma en el tiempo. En coherencia con lo anterior, se contemplan los siguientes:

1. Definición y evolución del ciclo de vida del dato en el DataR, incluyendo lineamientos de ingesta, almacenamiento, gestión de metadatos, retención, disposición (en concordancia con el ciclo de vida del dato) y organización de la información, en concordancia con las directrices definidas por la entidad.
2. Integración progresiva de nuevas fuentes de información al DataR, junto con su estructuración, alineación y mejora a los modelos de datos definidos y/o existentes.
3. Estructuración y mantenimiento de modelos de datos por dominio funcional, incluyendo su diseño, diccionarios de datos y mapeo de fuentes.
4. Definición e implementación de estándares para la construcción y gestión de flujos de datos, incluyendo integración, transformación, orquestación y manejo de grandes volúmenes de información.
5. Evolución y orquestación de esquemas de gobernanza de datos mediante el uso de Unity Catalog, operando bajo una arquitectura Medallón. Cabe resaltar que, actualmente DataR dispone de una estructura de organización y administración de datos segregados mediante catálogos dedicados para los entornos de Desarrollo (DEV), Pruebas (TEST) y Producción (PROD), y esta estructura permite una organización por dominios de negocio con una gestión de accesos centralizada, políticas de seguridad granulares y una administración de permisos activa en cada etapa del ciclo de vida del dato.
6. Implementación de estructuras para la gestión de calidad y trazabilidad de datos, incluyendo reglas de validación, control de inconsistencias y manejo de excepciones.
7. Definición y adopción de lineamientos para la integración y despliegue de productos de datos, incluyendo control de versiones y promoción entre ambientes.
8. Definición y actualización de la hoja de ruta y madurez del DataR, incluyendo la evolución de capacidades, la identificación de dependencias técnicas y consideraciones para la planificación de iniciativas conforme a las prioridades y dominios de negocio definidos por la entidad, así como la adopción de prácticas de gestión de datos y transferencia de conocimiento para su operación y mejora continua.

9. Definición de lineamientos para la optimización del uso de recursos de cómputo en los procesos de datos y documentación técnica de las estructuras, flujos y productos implementados.
10. Definición y adopción de esquemas de interoperabilidad de datos, incluyendo estándares de integración, intercambio y exposición de información entre sistemas internos y externos, en concordancia con los lineamientos técnicos de la entidad.
11. Definición e implementación de estrategias para la gestión de los cambios en las fuentes de datos (Change Data Capture - CDC), de acuerdo con las características de los dominios de información y las capacidades tecnológicas de cada fuente. La solución deberá contemplar tanto el uso de funcionalidades nativas de CDC disponibles en los sistemas origen como la implementación de mecanismos alternativos o desarrollos específicos cuando dichas capacidades no estén disponibles y puedan ser implementadas como un producto técnico.
12. Definición, documentación y acompañamiento técnico para la habilitación de capacidades tecnológicas requeridas por los productos implementados bajo la demanda que defina la entidad, aplicable únicamente a aquellos productos que esta determine. Esto incluye la identificación de necesidades de configuración, parametrización, accesos, seguridad, conectividad, despliegue, rendimiento, automatización e integración sobre los componentes de la arquitectura DataR. El proveedor deberá elaborar las especificaciones técnicas, recomendaciones, guías de implementación, criterios de validación y brindar el acompañamiento necesario para su correcta adopción, puesta en producción y operación.

Es importante precisar que el proveedor del servicio seleccionado no tendrá responsabilidad sobre la operación de la infraestructura, las redes, ni la administración del software base, así como tampoco tendrán obligaciones sobre lo concerniente a la operación, administración, definición y/o establecimiento de lineamientos de gobierno de datos. Estas tareas o responsabilidades permanecerán a cargo de las áreas correspondientes al interior de la entidad. No obstante, a lo anterior, los proponentes deberán conocer, seguir e implementar todas las políticas y lineamientos de la entidad en los productos construidos.

Las actividades se desarrollarán sobre la arquitectura tecnológica definida por la Entidad, en la cual se deben integrar los procesos de ingesta, transformación, almacenamiento, modelado y explotación de datos. Esta arquitectura soporta tanto la gestión de la información como su aprovechamiento para el consumo general al interior de la entidad.

5.2 Enfoque técnico a considerar para la prestación de servicios

El servicio requerido se orienta a la ejecución de procesos de ingeniería de datos sobre el DataR, abordando de manera integral el ciclo de vida del dato desde su integración hasta su consumo mediante la construcción, implementación, evolución y mantenimiento de los productos de

Información. En este contexto el/los proveedores deberán ejecutar actividades técnicas en marco a la disponibilización, integración, procesamiento, control y gestión de la información, en articulación con los lineamientos establecidos por la DIAN.

De manera específica, se deberán desarrollar los productos de información, teniendo en cuenta la ejecución de las siguientes actividades y/o entregables:

1. Realizar el levantamiento y documentación de los requerimientos funcionales asociados a la disponibilización y migración de datos, articulando las necesidades del negocio con los lineamientos técnicos del DataR, mediante la identificación de fuentes de información, dominios de datos, reglas de negocio, criterios de calidad, periodicidades de carga y condiciones de operación; así mismo, apoyar la definición del alcance técnico de los requerimientos y acompañar su ejecución en actividades de diseño, desarrollo, pruebas y despliegue, entre otras, verificando que los resultados obtenidos sean consistentes acorde con las condiciones definidas para su uso en los procesos la entidad.
2. Diseñar e implementar flujos de disponibilización de fuentes de datos en los ambientes de desarrollo, pruebas y producción para los ejercicios de migración, utilizando las herramientas proporcionadas por la entidad y permitiendo su correcta integración al DataR.
3. Construir mecanismos de seguimiento automatizado sobre los flujos y productos de información implementados, mediante la definición de métricas, el registro y notificación de fallos y errores, y la identificación de oportunidades de mejora en los procesos de datos.
4. Analizar, probar y validar la información resultante de la implementación de procesos ETL, conforme a los lineamientos de seguridad y buenas prácticas.
5. Identificar y definir reglas de validación de calidad de datos aplicables a un conjunto de tablas o conjuntos de información en el repositorio, considerando dimensiones como completitud, consistencia, integridad, disponibilidad, precisión y validez de los datos, en concordancia con estándares de la industria y lineamientos de la DIAN.
6. Implementar y gestionar los metadatos, el catálogo y el linaje de datos, incluyendo la definición de estructuras de información, trazabilidad de los datos y su relación con los procesos y fuentes de origen.
7. Diseñar e implementar mecanismos automatizados para la ejecución de validaciones de calidad de datos sobre conjuntos de información seleccionados, utilizando las herramientas tecnológicas disponibles en la arquitectura del DataR.
8. Generar métricas e indicadores de calidad y usabilidad los productos de información y de datos que permitan evaluar el cumplimiento de las reglas definidas y realizar seguimiento al comportamiento de la información procesada en el repositorio.

9. Construir instrumentos automatizados de seguimiento que permitan visualizar los resultados de las validaciones de calidad de datos y facilitar el seguimiento al estado de la información disponible en el repositorio.
10. Documentar los procesos desarrollados, incluyendo las reglas de calidad implementadas, los mecanismos de validación automatizados y las herramientas utilizadas, con el propósito de apoyar procesos de transferencia de conocimiento y la continuidad operativa.
11. Adoptar, fortalecer y gestionar procesos de integración y despliegue continuo (CI/CD), mediante la automatización de pipelines, la incorporación de controles de calidad, seguridad y versionamiento, y la administración de lo que se construya en marco a la implementación de los productos de información usables y productos técnicos en los ambientes de desarrollo, pruebas y producción, conforme a los lineamientos técnicos, de arquitectura y de operación definidos por la DIAN, facilitando la entrega continua, confiable y trazable de los componentes y productos desarrollados.

Elaborar la documentación técnica de los scripts, modelos de datos, productos y servicios desarrollados e implementados, incluyendo la documentación del resultado de las pruebas, en herramientas como Enterprise Architect versión 17 o superiores, o aquellas de índole institucional que puedan usarse para tal fin, conforme a los procedimientos y lineamientos internos definidos por la DIAN.

12. Ejecutar actividades de transferencia de conocimiento orientadas a la apropiación de los desarrollos y a la continuidad operativa de los productos y procesos implementados, esto mediante metodologías que le permitan a la entidad el fortalecimiento de sus capacidades técnicas. Estas actividades podrían incluir, entre otras, talleres prácticos, sesiones de mentoría técnica, esquemas de acompañamiento experto aprendiz, elaboración de guías y demás mecanismos que faciliten la adopción.
13. Implementar y definir estrategias de auditoría y trazabilidad sobre el acceso, uso y visibilidad de los productos de datos dependiendo del dominio y la sensibilidad de la información alineado con Unity Catalog.

Alienado con los numerales anteriores (5.1.1 y 5.1.2), el alcance entonces comprenderá la totalidad de las actividades técnicas, los productos técnicos y de información y las capacidades requeridas para la evolución, sostenibilidad y aprovechamiento del DataR, conforme a las necesidades de la Entidad y a los lineamientos técnicos, de arquitectura y de gestión de datos establecidos.

El cumplimiento del alcance se materializará en la entrega de productos técnicos y/o de información usables y escalables en el ambiente de producción del DataR, conforme a los criterios técnicos y funcionales definidos por la entidad, los cuales constituyen el resultado esperado del servicio.

La estimación del esfuerzo requerido para el desarrollo de los productos de información usables y/o técnicos descritos anteriormente, se debe enmarcar en los siguientes niveles de complejidad técnica:

Nivel de Complejidad	Definición
Baja	<p>Corresponde a desarrollos o ajustes de alcance acotado, orientados a la integración, transformación o disponibilización de datos mediante reglas de negocio simples y bajo impacto técnico.</p> <p>Incluye procesamiento sobre fuentes conocidas, estructuras estables y componentes reutilizables, con mínima dependencia entre procesos y baja afectación sobre la arquitectura existente.</p> <p>Generalmente contempla configuraciones estándar, transformaciones básicas, cargas incrementales simples, documentación técnica y validaciones elementales de calidad de datos.</p>
Media	<p>Corresponde a desarrollos e implementaciones que requieren integración de múltiples fuentes de información, aplicación de reglas de negocio intermedias, validaciones de calidad, procesamiento incremental y orquestación de flujos de datos.</p> <p>Implica interacción con diferentes componentes de la arquitectura DataR, manejo de volúmenes considerables de información, definición y ajuste de modelos de datos, automatización de procesos y ejecución de pruebas técnicas y funcionales. Puede requerir coordinación entre equipos técnicos y análisis de desempeño, trazabilidad y escalabilidad.</p>
Alta	<p>Corresponde a desarrollos e implementaciones estratégicas o de alta criticidad técnica y funcional asociados al procesamiento masivo de datos, múltiples integraciones y transformación compleja de información, así como a la construcción de capacidades transversales para el DataR.</p> <p>Incluye escenarios de alta volumetría, procesamiento distribuido, gestión de datos históricos, interoperabilidad entre dominios, automatización avanzada, gobierno y trazabilidad integral de datos, así como validaciones de calidad orientadas a garantizar la integridad del dato, disponibilidad y escalabilidad de las soluciones, e incorporación de análisis de costo-beneficio aplicando principios FinOps.</p> <p>Requiere análisis avanzado de arquitectura, optimización de recursos, coordinación técnica y gestión de riesgos debido al impacto operativo e institucional de los productos implementados.</p>

7. SOLICITUD DE INFORMACIÓN

Teniendo en cuenta la información anterior, se solicita a las firmas interesadas proporcionar la siguiente información, la cual será utilizada para comprender mejor el mercado y las soluciones disponibles y permitirá soportar y enriquecer el eventual proceso de adquisición. Por favor tener en cuenta:

- La información debe ser diligenciada, de la forma más específica posible, en los formatos proporcionados [Ej.: Planillas en Excel, Cuestionarios en Word].
- Cualquier pregunta o aclaración sobre este RFI debe realizarse por escrito al mismo correo electrónico, a más tardar en la fecha establecida como límite para la elaboración de preguntas por parte de los interesados. Las respuestas serán publicadas en la página web.
- El plazo límite para la entrega de los anexos diligenciados por parte de los interesados quedará establecida en el aviso de publicación en la página web.
- La información debe ser enviada al correo electrónico: adquisiciones@fondodian.gov.co que es gestionado por la Unidad de Coordinación del Programa de Apoyo a la Modernización de la DIAN.
- Todas las interacciones entre la DIAN y los interesados en participar en este requerimiento se deben realizar utilizando el correo mencionado. No se aceptarán respuestas al RFI que se entreguen por un medio diferente o que se entreguen en papel en las dependencias de la DIAN, indicando la siguiente denominación en el asunto del correo: **RFI – Servicios de Desarrollo Repositorios de Datos DATAR.**
- La DIAN se reserva el derecho de analizar las respuestas entregadas por los interesados y de solicitar las aclaraciones que a su juicio se requieran, sin que ello implique un compromiso de contratación posterior.

6.1 Información General del Interesado

El **Anexo 1 – Información Interesado RFI DATAR.xlsx** tiene como propósito recopilar la información general del interesado en participar en el RFI de servicios para DataR DIAN. La información solicitada permitirá conocer la identificación de la firma, su experiencia, capacidades técnicas, presencia en Colombia, certificaciones, alianzas estratégicas, metodologías de trabajo y soluciones o servicios relacionados con plataformas de datos.

Estos insumos son relevantes para la etapa de análisis de mercado y estructuración de un eventual proceso contractual, permitiendo identificar las capacidades disponibles en el mercado, las condiciones bajo las cuales podrían prestarse los servicios requeridos y las recomendaciones que los interesados consideren pertinentes para fortalecer el alcance, enfoque y ejecución del objeto requerido por la DIAN.

6.2 Capacidades Técnicas y Experiencia

Se requiere que la firma interesada diligencie la experiencia con la que cuenta, a fecha en contratos de desarrollo, implementación, operación o mantenimiento de soluciones similares relacionados con el alcance de este RFI. Para ello, por favor diligencie la información en el **Anexo 2 – Capacidades y experiencia del Interesado RFI DATAR.xlsx**

En cuanto a la experiencia específica, se desea conocer lo referente a:

1. Grandes volúmenes de datos (escalas Terabytes (TB) / Petabytes (PB)).
2. Entornos con alta disponibilidad, continuidad operativa, tolerancia a fallos, recuperación ante desastres, monitoreo, soporte, gobierno, seguridad y calidad.
3. Desarrollo de Código Resiliente: código robusto, que incluya el manejo de excepciones, mecanismos de reintento y logging avanzado en proyectos de datos críticos.
4. Gestión del Reúso de Código: estándares de codificación y patrones de diseño que promuevan el reúso efectivo y mantenible de componentes de código en entornos de Big Data.
5. Diseño e implementación (D&I) de arquitecturas de gobierno de datos, analítica avanzada y soluciones de inteligencia artificial.
6. Desarrollo e implementaciones (D&I) con estándares o lineamientos de calidad técnica, adoptando procesos de mejora continua.

De igual forma, se desea conocer la , los mecanismos que ha utilizado para garantizar la continuidad de su servicio, la estabilidad su equipo de trabajo y la adecuada transferencia de conocimiento durante la ejecución contractual

6.3 Características de las Soluciones Tecnológicas Propuestas

El **Anexo 3 – Características Solución RFI DATAR.xlsx** tiene como propósito recopilar toda la información respecto a las capacidades técnicas que ha desarrollado e implementado la firma interesada de acuerdo con la experiencia que tenga en cada uno de sus proyectos/contratos de desarrollo, implementación, operación o mantenimiento de soluciones similares relacionados con el objeto y alcance de este RFI.

La información solicitada permitirá identificar las capacidades técnicas disponibles en el mercado para atender este tipo de servicios de desarrollo de productos, las condiciones bajo las cuales se pueden prestar en forma eficiente y las recomendaciones que los interesados consideren pertinentes para fortalecer el alcance de estos.

Se solicita diligenciar en el anexo toda la información respecto a:

- Estrategias, prácticas y herramientas para gestionar la continuidad operativa, la mejora continua y la gestión eficiente del DataR, el marco operativo (DataOps), el manejo de

errores, excepciones y resiliencia en pipelines y productos de datos, el reúso de código en el desarrollo y mantenimiento de soluciones de datos, el modelo de 'Data-as-a-Product', la continuidad del servicio, el modelo de exposición y consumo de datos, pruebas de interoperabilidad, la integridad semántica de entidades / atributos y la alineación con modelos canónicos desde DataR, las capacidades de metadatos, catálogo, linaje y trazabilidad de datos, los ANS, la gestión de riesgos.

- En cuanto a la gestión de seguridad de la información, se desea conocer las capacidades, experiencias y prácticas implementadas por los interesados en materia de seguridad de la información, protección de datos, cumplimiento normativo y continuidad operativa.
- En cuanto a la innovación del servicio y el valor agregado, es importante conocer los enfoques, prácticas y/o tecnologías relevantes que permitan imprimirle valor de innovación a los productos técnicos y/o de información objeto de este RFI.

6.4 Equipo Base del Proveedor

La DIAN ha identificado un equipo base que podría atender la ejecución de los servicios contemplados en este documento. Sin embargo, es importante obtener, desde su experiencia, la información de los roles y sus características para atender este tipo de soluciones y garantizar así la ejecución en tiempo y forma de los diferentes productos de información requeridos. Para ello, se requiere diligenciar el **Anexo 4 – Información equipo de trabajo RFI DATAR.xlsx**

6.5 Aspectos Financieros (Estimaciones Preliminares)

De acuerdo con la información proporcionada en este documento, se requiere diligenciar en el **Anexo 5 – Cuadro de Precios RFI DATAR.xlsx** el valor hora por los servicios de desarrollo de los requerimientos técnicos y/o de información, clasificados según su nivel de complejidad, acorde con su experiencia en el mercado de los productos técnicos y de información descritos para este RFI.

Para facilitar el diligenciamiento, se ofrecen unas cifras de referencia (y en ningún caso proyectadas o comprometidas por la DIAN) respecto a la cantidad de horas que se demandarán por cada complejidad. Estas cifras fueron estimadas a nivel interno de acuerdo con las experiencias técnicas previas y el juicio de expertos de la DIAN.

Nivel de Complejidad	Horas de Referencia (Estimadas)
Baja	800 horas
Media	1.621 horas
Alta	4.000 horas

También es relevante conocer, a nivel de riesgos de ejecución presupuestal, qué elementos han sido considerados dentro de la experiencia con la que cuenta la firma interesada.

Es importante aclarar que el modelo previsto para un eventual proceso de contratación correspondería a un esquema de consumo abierto a demanda. Por lo tanto, la prestación efectiva de los servicios estará sujeta a las necesidades que determine la DIAN durante la etapa de ejecución, sin que ello obligue un consumo mínimo.