ANEXO N°1
ESPECIFICACIONES TÉCNICAS

"Desarrollo de una solución tecnológica para solicitar y generar reportes de forma asíncrona desde datos en medallones (delta lake) de diversos documentos electrónicos procesados por la DIAN"

1. Objetivo

Desarrollar una aplicación no corporativa que permita a los usuarios internos (funcionarios) de la entidad solicitar reportes predefinidos de documentos electrónicos procesados por la DIAN, así como los ETL's necesarios para generar los resultados de forma asíncrona aprovechando los datos disponibles en Databricks, para facilitar el acceso a la información de la entidad.

2. Alcance

El alcance del proyecto comprende 2 componentes:

1. El desarrollo de una aplicación ("solución no corporativa") mediante el módulo "APP Databricks" en los recursos disponibles de espacios de trabajo databricks de la DIAN, para registrar las solicitudes de reportes.

2. El desarrollo de los ETL's necesarias para atender las solicitudes de forma asíncrona (que en conjunto con las demás solicitudes que se acumulen en una hora se procesarán por lotes hasta generar los resultados esperados en tablas delta del medallón gold y reportes materializados como archivos html y csv en volúmenes disponibles dentro de unity catalog)

El proyecto se ejecutará bajo un modelo de desarrollo, donde el contratista trabajará según los lineamientos de los equipos de la DIAN para entregar una solución que cumpla con las necesidades de negocio y los estándares técnicos de la entidad.

El alcance del desarrollo se estructura en los siguientes pilares:

- **Fundación de Datos Analíticos en Capa Oro ("Gold"):**
    - Diseñar los conjuntos de datos agregados (resultados procesados) que residirán en la capa Gold de la arquitectura medallón, de acuerdo con las reglas de negocio, requisitos funcionales y no funcionales proporcionados por la DIAN.
    - Desarrollar, integrar y desplegar los pipelines de datos (ETL/ELT) en los entornos definidos por la DIAN. Estos procesos, basados en notebooks de PySpark/SQL, poblarán las tablas agregadas. El contratista ejecutará el despliegue siguiendo los lineamientos del equipo técnico de la DIAN, quien validará el cumplimiento de los estándares de arquitectura y seguridad.
    - Generar los pull request para obtener la aprobación para las promociones de código entre ambientes.
    - La construcción de estos modelos de datos agregados es un pilar fundamental del proyecto. El objetivo es pre-calcular indicadores y dimensiones que permitan a la

capa de presentación responder de manera óptima a consultas complejas, como el filtrado por uno o múltiples NITs, rangos de fechas, y categorías de negocio. La definición de estas métricas, los tipos de gráficos y las visualizaciones específicas para cada reporte se realizará en las mesas de trabajo lideradas por el contratista, dentro del marco definido en los criterios de aceptación.

- **Interfaz de registro (Capa de Presentación):**
    - El equipo contratista desarrollará una **Aplicación de Solicitudes** (Databricks App) que servirá como interfaz principal para que los usuarios de negocio registren, parametricen y monitoreen el estado de sus solicitudes de reportes.
    - El equipo contratista asegurará que los productos que implemente se traducen tanto en la materialización de informes detallados a través de archivos html estandarizados para cada tipo de reporte definido como en archivos csv, y esos elementos serán consistentes con los registros en las tablas del medallón oro.
- **Solución de Procesamiento y Generación de resultados (Capa de Procesamiento):**
    - El equipo contratista implementará ETL's que procesen las solicitudes registradas para actualizar su estado y finalmente, generar resultados en tablas delta, archivos html y csv.
    - El equipo contratista garantizará que todos los reportes generados (solicitados desde la aplicación para registrar la solicitud) se puedan descargar en formato CSV y HTML desde los volúmenes en unity catalog dispuestos a las áreas de negocio.
    - El equipo contratista deberá garantizar que los resultados generados se pueden consultar desde las tablas que se generen en medallón oro.
- **Gobernanza, Calidad y Seguridad:**
    - El equipo contratista implementará según las definiciones de las áreas de negocio, las reglas de calidad de datos sobre los pipelines y modelos de la capa Gold para asegurar la confiabilidad de los informes.
    - Los encargados de QA en el equipo contratista definirán los criterios y objetivos de calidad a fin de garantizar que los productos generados cumplan con las expectativas de cliente y legislación vigente.
    - Los encargados de QA en el equipo contratista verificarán y validarán los resultados generados de manera técnica a fin de identificar y reportar los defectos que se encuentren de forma que puedan corregirlos antes de solicitar alguna promoción de código fuente.
    - El equipo contratista seguirá los lineamientos o reglas de seguridad que el área técnicas de la DIAN haya definido, de igual forma el equipo técnico se encargará de limitar los distintos accesos y permisos exclusivamente a quienes deben tenerlos a través de unity catalog.

- El equipo contratista implementará los productos en los ambientes de trabajo y herramientas dispuestas por las áreas técnicas de la Entidad.
- El equipo contratista diseñará los modelos de datos y las visualizaciones para que sean compatibles con las políticas de seguridad de Unity Catalog, permitiendo la implementación de control de acceso y seguridad a nivel de fila (Row-Level Security), para garantizar que cada área de negocio solo acceda a la información que le compete, para ello tendrá que identificar los elementos necesarios en cada conjunto que se genere que permitan aplicar tal enfoque.

3. Definiciones, siglas y abreviaturas

Término	Definición
ACID	Acrónimo de **A**tomicidad, **C**onsistencia, **A**islamiento y **D**urabilidad. Conjunto de propiedades que garantizan la fiabilidad e integridad de las operaciones de datos. En el proyecto, Delta Lake implementa estas garantías sobre el Data Lake.
ADB (Azure Databricks)	Plataforma de análisis unificada y abierta para crear, implementar, compartir y mantener soluciones de datos, análisis e IA a escala, integrada con la nube de Azure.
ADLS Gen2	Azure Data Lake Storage Gen2. Servicio de almacenamiento en la nube de Microsoft optimizado para cargas de trabajo de análisis de Big Data, que sirve como base para el Lakehouse de Databricks.
App	Abreviatura de aplicación digital, usualmente una aplicación web o móvil diseñada para una función específica. En este anexo, se refiere a una "Databricks App".
Arquitectura Medallón	Enfoque de diseño de datos en un Lakehouse que organiza los datos en tres capas según la evolución de su calidad: **Bronce** (datos “crudos” sin transformaciones o exclusivamente las mínimas necesarias para permitir los demás pasos en el proceso, por ejemplo generar esquema), **Plata** (limpios, validados y certificados por dueño de datos) y **Oro** (basados en datos previamente certificados, para ser enriquecidos, agregados y listos para el uso por áreas de negocio y operaciones analíticas).
CI/CD	Acrónimo de Integración Continua (Continuous Integration) y Despliegue Continuo (Continuous Deployment). Prácticas de DevOps para automatizar las fases de construcción, prueba y despliegue de software.
Databricks App	Aplicación web que se ejecuta como un servicio en contenedores en la plataforma sin servidor de Databricks, construida con frameworks como Streamlit o Dash.

Delta Lake	Capa de almacenamiento de código abierto que aporta fiabilidad, calidad y rendimiento a los Data Lakes, habilitando transacciones ACID y otras funcionalidades avanzadas sobre archivos Parquet.
ETL/ELT	Acrónimos para procesos de datos. **E**xtraer, **T**ransformar y **C**argar (ETL) y **E**xtraer, **C**argar y **T**ransformar (ELT). Se refiere a los pipelines que mueven y procesan datos entre sistemas.
Framework	Marco de trabajo de software que proporciona una estructura base para desarrollar aplicaciones, como Streamlit, Dash o Gradio.
Lakehouse	Arquitectura de datos moderna que combina la flexibilidad y bajo costo de un Data Lake con la fiabilidad y el rendimiento de un Data Warehouse, implementada en este caso sobre Databricks y Delta Lake.
Microsoft Entra ID	Servicio de gestión de identidades y accesos basado en la nube de Microsoft, utilizado por la DIAN para la autenticación de usuarios.
MVP (Minimum Viable Product)	Versión inicial de un producto que incluye solo las características esenciales para ser funcional y satisfacer a los primeros usuarios, permitiendo recoger feedback para futuras iteraciones.
Notebook Interactivo	Herramienta principal para el desarrollo en Databricks, que permite combinar código (ej. PySpark, SQL), texto y visualizaciones en un único documento colaborativo.
OLTP (Procesamiento de Transacciones en Línea)	Tipo de sistema informático que gestiona aplicaciones orientadas a transacciones, como los sistemas de facturación en tiempo real. Este proyecto **no** es un sistema OLTP.
Pipeline de Datos (también llamado “canalización”)	Conjunto de procesos automatizados que mueven y transforman datos desde un origen a un destino. En este proyecto, se refiere a los “Jobs” de Databricks que construyen la capa “Gold”.
PMV	Siglas para Producto Mínimo Viable (*en idioma inglés las siglas son MVP por Minimum Viable Product*)
PySpark	API de Python para Apache Spark, que permite a los desarrolladores escribir aplicaciones de análisis de Big Data utilizando el lenguaje de programación Python.
QA (Aseguramiento de Calidad)	Sigla en inglés para *Quality Assurance*. Proceso o conjunto de procesos utilizados para medir y asegurar la calidad de un producto.
RBAC (Control de Acceso Basado en Roles)	Sigla en inglés para *Role-Based Access Control*. Modelo de seguridad que restringe el acceso al sistema a los usuarios autorizados basándose en los roles que se les asignan.

Reporte	Medio de comunicación que proporciona una visión específica y objetiva de información relevante sobre un asunto, presentando datos de manera sistemática y estructurada.
Row-Level Security (RLS)	Seguridad a Nivel de Fila. Funcionalidad de seguridad que restringe el acceso a las filas de una tabla de datos basándose en la identidad o el rol del usuario que realiza la consulta.
SFESO	Subdirección de Factura Electrónica y Soluciones Operativas de la DIAN.
SPD	Subdirección de Procesamiento de Datos de la DIAN.
UAT (Pruebas de Aceptación de Usuario)	Siga en inglés para “User Acceptance Testing”. Fase final del proceso de pruebas de software en la que los usuarios finales validan que el sistema cumple con sus necesidades de negocio en escenarios del mundo real.
UI (Interfaz de Usuario)	Sigla en inglés para *User Interface*. Medio a través del cual un usuario interactúa con una máquina, dispositivo o programa informático.
Unity Catalog	Solución de gobernanza unificada para datos y activos de IA en la plataforma Databricks, que proporciona control de acceso centralizado, auditoría y linaje de datos.
UX (User Experience)	Siglas de User Experience o Experiencia de Usuario, se refiere a las percepciones, emociones y respuestas que una persona tiene al interactuar con un producto, sistema o servicio digital. El objetivo del diseño UX es asegurar que esta interacción sea útil, usable, agradable y, sobre todo, que cumpla con las necesidades y objetivos del usuario

4. Descripción de interesados

Nombre/rol	Funciones y responsabilidades
Equipos funcionales	• **Usuarios SFESO** (solicitantes de la aplicación, principales responsables de los conjuntos de datos de documentos electrónicos).
	• **Usuarios de negocio de áreas distintas a la SFESO** (dueños de otros conjuntos de datos como por ejemplo RUT).
	• **Usuarios finales** de la solución (consumidores de la aplicación, que pueden estar en cualquier área de la DIAN, ya sean usuarios sin conocimiento técnico que sólo usarán el archivo html, usuarios que manejan excel que usarán los csv, o usuarios que pueden tener conocimientos técnicos mínimos en sql y python).

Equipo Contratista	• **Diseñar, desarrollar, integrar y solicitar los recursos necesarios (permisos, configuración y la promoción de productos /despliegue) de la solución tecnológica completa**, incluyendo: ➢ Las definiciones de tablas, nomenclaturas y características (esquemas) que serán pobladas en medallón oro dentro de unity catalog al atender automáticamente las solicitudes de reportes tipificados. ➢ Los pipelines de datos (ETL/ELT) para agregar a los conjuntos nuevos que en capa “oro” representarán las atenciones a solicitudes resueltas. ➢ La aplicación de registro de solicitudes de reportes (Databricks App). ➢ Las plantillas html para visualización de reportes para usuarios sin conocimientos técnicos que definen los contenidos a materializar en html mediante ETL’s ➢ Plantillas de notebooks para visualización. • **Liderar las mesas de trabajo** para definir y documentar las especificaciones funcionales y técnicas de cada reporte, en conjunto con las áreas funcionales y técnicas de la DIAN. • **Ejecutar el despliegue técnico** de todos los componentes en los entornos de la DIAN desarrollo, pruebas y producción), siguiendo los lineamientos y tras aprobaciones del equipo técnico de la DIAN. • **Implementar las reglas de calidad y seguridad** de datos definidas con el equipo funcional y el equipo técnico DIAN. • **Implementar las reglas de seguridad** de datos definidas con el equipo técnico DIAN. • **Generar toda la documentación del proyecto**: manuales técnicos, manuales de usuario, actas de prueba y formatos de entrega requeridos por la entidad. Son los responsables de: • La implementación de productos • La implementación de estrategias para generar productos y asegurar calidad • Asegurar la calidad de productos, datos y lograr la satisfacción de las áreas de negocio DIAN para lograr la aprobación de los elementos implementados • Documentación y transferencia de conocimiento • Garantía del producto posterior a su implementación • Soporte técnico del producto dentro de su garantía
Equipo Técnico DIAN	Conformado por los equipos designados desde las subdirecciones: • SIA • SPD Adicionalmente, pueden solicitar apoyo de otros equipos en la dirección de innovación y tecnología como lo son: • SSD (subdirección de soluciones y desarrollo) • STO (subdirección infraestructura tecnológica y de operaciones)

	Y trabajan en conjunto o solicitan labores específicas mediante procedimientos internos a proveedores externos dentro del proyecto *Multinube híbrida*
SFESO	**Subdirección de Factura Electrónica y soluciones operativas**, responsable de definir, documentar y validar las reglas de negocio aplicables a todos los documentos electrónicos procesados por la DIAN. Dentro de sus labores para este proyecto deben: • Actuar como experto funcional en la normatividad, estructura y semántica de los datos de origen, siendo el interlocutor clave para resolver cualquier duda sobre su interpretación. • Validar que la implementación realizada por el contratista en los reportes y pipelines refleje fielmente la normativa y las reglas de negocio establecidas. • Definir la lógica de negocio para los reportes. • Validar la lógica de negocio de los reportes. • Principales expertos por parte de la DIAN que trabajarán con el contratista en la definición de requerimientos, reglas de cálculo y criterios de aceptación para cada informe en los conjuntos de los que son expertos (documentos electrónicos). • Colaborador con otras áreas de negocio expertas en la DIAN con quienes trabajarán junto con el contratista en la definición de requerimientos, reglas de cálculo y criterios de aceptación para cada informe en los conjuntos distintos a documentos electrónicos que puedan requerirse. • Realizar las pruebas de aceptación de usuario (UAT) para validar que los resultados cumplen con las necesidades operativas y estratégicas.
SIA	Subdirección de Información y Analítica Incluye a los equipos de: • Analítica • Gobernanza Subdirección responsable de apoyar a la SFESO en labores que le sean solicitadas relacionadas a validar conjuntos generados de forma interna para contrastar con las labores que realice el equipo contratista. También podrán aportar al equipo de negocio con su conocimiento en BI para la evaluación de productos materializados como html, a fin de corroborar las labores relacionadas a validaciones de tal producto que el equipo contratista entregue como evidencia.
SPD	Subdirección procesamiento de datos, incluye los equipos internos: • Arquitectura de datos • Solicitudes de información • Administración de datos Encargada de indicar al equipo contratista los lineamientos de:

	• Nombrado de elementos como tablas, catálogos y otros componentes en unity catalog • Solicitud de permisos Responsables de: • Compartir los lineamientos o políticas existentes relacionados a seguridad y uso de las herramientas disponibles en el ecosistema Databricks • Definir y proveer los lineamientos generales técnicos, arquitectónicos y de seguridad que el contratista debe seguir. • Administrar el Unity Catalog de Databricks, la gestión de permisos y la infraestructura subyacente, dentro de tales labores se incluye: o Brindar permisos en unity catalog o Aprobar los pull request en azure DevOps *(Aceptar los componentes desarrollados como modelos de datos traducidos a tablas, etl's, entre otros elementos… antes de su promoción a producción).* o Apoyar a los equipos de negocio a petición explícita cuando tengan inconvenientes relacionados a permisos en los conjuntos y espacios de trabajo. o Garantizar el acceso del contratista a los entornos, datos y herramientas necesarias para el desarrollo. o Revisar las evidencias que aporte el líder técnico del equipo contratista aporte tras el *code review* que realizó para los desarrollos previos a despliegues, y realizar una nueva revisión como lectura de validación final en el código (directamente en azure DevOps) para identificar cualquier incompatibilidad con los estándares de la DIAN previo a aprobar despliegue. Para estas labores podrán solicitar espacios con el líder técnico del equipo contratista quien es el encargado de la calidad del código que genere su equipo. • Analizar necesidad de creación de nuevos recursos cloud (espacios de trabajo, cuentas de almacenamiento, contenedores, etc) y Solicitar creación o modificación de recursos al responsable contractual del proyecto Multinube híbrida relacionados al repositorio único de datos (DataR). • Actuar como custodio de la arquitectura de datos global, asegurando que la solución es compatible con el ecosistema DataR.

5. Equipo Mínimo Requerido por el Contratista

Para garantizar la correcta ejecución del proyecto en tiempo y forma, el contratista deberá proponer un equipo de trabajo que cuente, como mínimo, con los siguientes perfiles y capacidades demostrables:

Perfil Mínimo	Responsabilidades Clave en el Proyecto	Formación Académica	Requisitos y Experiencia Mínima Exigida
1. Líder del equipo contratista	Liderar el equipo contratista y facilitar las mesas de trabajo, asegurando que los requerimientos de negocio se traduzcan en historias de usuario con criterios de aceptación de datos claros y medibles. Gestionar proactivamente los riesgos inherentes a proyectos de datos, como la calidad de los datos de origen, la complejidad de las reglas de negocio y las dependencias entre equipos. Liderar las gestiones internas de equipo contratista y ser puente de comunicación con los líderes de equipos DIAN que definen y aceptan productos, brindan permisos a recursos o confirman ejecuciones de procesos según las especificaciones/requisitos.	Título profesional en Ingeniería de Sistemas, o Ingeniería de software, o Diseño Industrial, o Telecomunicaciones, o Eléctrica, o Electrónica, o Industrial o los núcleos básicos de conocimiento de esas profesiones de acuerdo con lo establecido en el SNIES. Título de posgrado en alguna de las siguientes áreas: Gestión de Proyectos o Gerencia de proyectos informáticos o Gerencia de proyectos de sistemas o Cooperación internacional y gestión de proyectos para el desarrollo o Dirección de proyectos. Certificacion PMP o Scrum Master (CSM/PSM)) o similar.	• Experiencia general mínima de 5 años. • Mínimo 3 años de experiencia relacionada demostrable en proyectos de **ingeniería de datos, Business Intelligence o Analítica Avanzada**. • Experiencia comprobada en la gestión de **stakeholders de negocio no técnicos** para la definición de requerimientos analíticos.
2. Líder Técnico	Responsable del diseño técnico global y la integridad de las tablas, código implementado y elementos técnicos de los productos, desde la ingesta hasta la capa de consumo. Definir y documentar los patrones, los estándares de codificación y las mejores	Título de pregrado en las áreas del conocimiento de Ingeniería de Sistemas, o Software, Electrónica, o Telecomunicaciones, o Industrial o Mecatrónica o los núcleos básicos de	• Experiencia general mínima de 5 años • Mínimo 3 **años de experiencia** relacionada con el diseño e implementación de soluciones de datos en la nube Azure. • Experiencia comprobada en el diseño y optimización

	prácticas de rendimiento y optimización de costos en Databricks alineado con las herramientas y arquitecturas con las que cuenta la DIAN. Liderar la selección y aplicación de técnicas de modelado de datos (ej. Dimensional, Data Vault) para la capa Gold, asegurando que el diseño sea escalable, eficiente y responda a las necesidades de negocio. Ser quien se responsabiliza por la calidad técnica y el rendimiento de la solución. Realizar “code review” de los notebooks o archivos de código fuente que produce el equipo contratista.	conocimiento de esas profesiones de acuerdo con lo establecido en el SNIES. Con mínimo especialización en áreas de Ingeniería de sistemas o Ingeniería de Software o Arquitectura de Software o Desarrollo de Software o Construcción de Software o Gerencia Informática o en Big Data / inteligencia de negocios. Certificación DAMA y/o Certificación "Databricks Certified Data Engineer Asociate" y/o Certificacion Azure Solutions Architect Expert y/o Microsoft Implementing Data Engineering Solution y/o Azure Databricks Platform Architect y/o Microsoft Certified: Azure Data Engineer Associate (DP-203).	de **modelos dimensionales** para Data Warehouses o Lakehouses a gran escala. • Habilidades de comunicación para explicar decisiones de diseño técnico a funcionarios de la DIAN.
3. Ingeniero de requisitos (requerimientos)	Responsable de documentar todas las especificaciones funcionales mediante historias de usuario según la estrategia de objetivos “SMART” Ser quien se responsabiliza por la calidad de los documentos de especificación con las necesidades y reglas de las áreas de negocio,	Título de pregrado en las áreas del conocimiento de Ingeniería de Sistemas, o Electrónica, o Telecomunicaciones, o Industrial o Mecatrónica o los núcleos básicos de conocimiento de esas profesiones de acuerdo con lo	• Experiencia general mínima de 5 años • Mínimo 3 **años de experiencia** relacionada con especificación de requisitos mediante historias de usuario aplicando estrategia SMART

	incluyendo definiciones concisas y elementos medibles a implementar con criterios de aceptación que puedan ser ejecutadas por las áreas de negocio.	establecido en el SNIES.	
4. Dos (2) Ingeniero(s) de Datos	Desarrollar, probar y solicitan despliegues/promociones de los **pipelines de datos (ETL) robustos, monitoreables y eficientes** para construir y mantener los resultados procesados (reportes atendidos) en capa Gold. Implementar las transformaciones, agregaciones y reglas de negocio complejas en notebooks de **PySpark/SQL**, asegurando la trazabilidad y el linaje de los datos. Integrar **lógicas de calidad de datos y pruebas unitarias/integración** dentro de los pipelines para garantizar la fiabilidad de los datos.	Profesional en Ingeniería (Sistemas, o electrónica, eléctrica, o telecomunicaciones, o industrial o ***Administrador de Sistemas*** informáticos) o los núcleos básicos de conocimiento de esas profesiones de acuerdo con lo establecido en el SNIES.	• Experiencia Mínima 3 **años de experiencia** en ingeniería de datos en ecosistemas de Big Data. • Dominio intermedio o avanzado de **PySpark** y **SQL**, incluyendo funciones de ventana, optimización de joins y manipulación de estructuras de datos complejas. • Recomendable Experiencia con herramientas de control de versiones **(GIT)** y CI/CD
5. Dos (2) Desarrolladores python –	Diseñar y construir la Databricks App que permitirá registrar solicitudes de los tipos de reportes. Traducir los requerimientos de negocio en especificaciones técnicas para viabilizar los flujos de datos (canalizaciones, procesos en lotes) que procesen las solicitudes hasta generar reportes (materializados en archivos csv y html). Validar los datos desde una perspectiva funcional, realizando análisis exploratorios para identificar patrones, anomalías y asegurar que los resultados	Profesional en Ingeniería (Sistemas, o electrónica, eléctrica, o telecomunicaciones, o industrial o ***Ingeniero Informático o Ingeniería en Automatización industria***l) o los núcleos básicos de conocimiento de esas profesiones de acuerdo con lo establecido en el SNIES.	• Mínimo 1, recomendable **3 años de experiencia en** aplicaciones de Python, Node.js. Entre los marcos populares de Python recomendados se incluyen Streamlit, Dash y Gradio. también se admiten marcos de Node.js como React, Angular, Svelte y Express. (Para que pueda implementar una app databricks que requiere conocimiento en alguno de esos marcos recomendados). • Fuerte dominio de PYSPARK/SQL para análisis de datos complejos. • Capacidad para comunicar

	de los reportes sean coherentes y lógicos. Colaborar estrechamente con el Desarrollador web para asegurar que se hayan definido los tipos de reportes con suficiente refinamiento y el resultado se genere correctamente.		hallazgos y conocimientos derivados de los datos a audiencias no técnicas.
6. Desarrollador web enfocado en la visualización de datos mediante reportes html	Diseñar y construir los reportes html estáticos plantilla que sirvan como base para la generación de reportes automatizada mediante ETL's que implementa(n) ingeniero(s) de datos. Traducir los requerimientos de negocio en **especificaciones técnicas para los reportes**, definiendo métricas, KPIs y dimensiones. **Validar los datos desde una perspectiva funcional**, realizando análisis exploratorios para identificar patrones, anomalías y asegurar que los resultados de los reportes sean coherentes y lógicos. Colaborar estrechamente con el Desarrollador Frontend para asegurar que la experiencia de usuario en la aplicación sea analíticamente intuitiva.	Profesional en Ingeniería (Sistemas, o electrónica, eléctrica, o telecomunicaciones, o industrial ***o Ingeniero Informático o Ingeniería en Automatización industrial***), o los núcleos básicos de conocimiento de ingenierías de acuerdo con lo establecido en el SNIES.	• Mínimo 1, recomendable **3 años de experiencia en** informes html interactivos javascript (html estático basado en conjuntos de datos json) • Fuerte dominio de alguna de las siguientes librerías javascript: apache echarts, d3.js para generar reportes en html estático (1 archivo html por reporte) con visualizaciones que ayuden a comprender los resultados en el reporte. • **Capacidad para comunicar hallazgos y conocimientos** derivados de los datos a audiencias no técnicas.
7. Experto en UI/UX	Diseñar mockups para definir las pantallas de la databricks app Validar que la experiencia de usuario y la interfaz sean intuitivas y permitan a los solicitantes de reportes tener claridad sobre su solicitud, estado	Profesional en Ingeniería (Sistemas, o electrónica, eléctrica, o telecomunicaciones, o industrial) o afines a la Ingeniería (por ejemplo: diseño industrial, diseño	• Mínimo **3 años de experiencia** en roles de UX • Dominio avanzado de herramientas para mockups. • Experiencia en la **prueba y validación** de la experiencia de interfaces propuestas.

	Diseñar y desarrollar pruebas automáticas de integración y carga (stress) si se requieren. Documentar y gestionar las novedades, inconsistencias, hallazgos en conjuntos de datos, trabajando con los Ingenieros de Datos para su resolución y validación.	gráfico, etc, con énfasis en UX).	• Responsable de garantizar que las pantallas que se diseñen sean usables e intuitivas, orientadas a cumplir con los requisitos definidos manteniendo la imagen institucional y la facilidad de uso (requerir la menor cantidad de documentación o ayudas adicionales para usar las pantallas de solicitud) • Responsable de garantizar que un usuario puede encontrar e interpretar los reportes html *sin necesidad de tener conocimientos técnicos como sql o python*.
8. Ingeniero(s) de Calidad (QA) / Tester(s)	Diseñar y ejecutar un **plan de pruebas integral enfocado en la validación de datos**. Desarrollar y automatizar **casos de prueba para validar la lógica de los pipelines**, la correcta aplicación de reglas de negocio y la precisión numérica de los agregados. Realizar **validaciones de datos** entre la capa Gold y las capas de origen (Silver/Bronze) para asegurar que no haya pérdida o corrupción de información. Diseñar y desarrollar pruebas automáticas de integración y carga (stress). Documentar incidentes, problemas, inconsistencias, hallazgos en conjuntos de datos y novedades presentadas y gestionarlas, colaborando con los Ingenieros de Datos del	Profesional en Ingeniería (Sistemas, o electrónica, eléctrica, o telecomunicaciones, o industrial) o afines a la Ingeniería. Recomendable certificaciones, por ejemplo: ISTQB CTFL.	• Mínimo **3 años de experiencia** en roles de QA o Testing, con un enfoque específico en **proyectos de ETL, Data Warehouse** • Dominio avanzado de PYSPARK/SQL para escribir consultas de validación complejas y realizar perfiles de datos, manejo de herramientas como pytest. • Experiencia en la **prueba y validación de reglas de negocio en flujos de datos**. • Responsable de garantizar que los conjuntos tienen la calidad necesaria tanto a nivel técnico como funcional.

	equipo contratista para su resolución y validación. Acompañamiento a los usuarios funcionales durante la ejecución de las pruebas de aceptación. Corrección de bugs (comportamientos o características no planeados o distintos a los especificados para algún componente, funcionalidad o conjunto) reportados por los usuarios funcionales. Elaboración de los informes de pruebas		

Nota: El contratista deberá presentar las hojas de vida y las certificaciones del equipo propuesto para validar el cumplimiento de estos requisitos. La combinación de roles es aceptable siempre que un solo recurso cumpla con la totalidad de los requisitos exigidos para cada rol que asume.

6. Descripción Técnica del Producto

La solución se concibe como un ecosistema de componentes nativos de la nube de Azure y la Plataforma Databricks, diseñado para operar de forma desacoplada, segura y escalable. La arquitectura se fundamenta en los siguientes pilares tecnológicos y de diseño:

6.1. Arquitectura General y Flujo de Datos

El sistema opera bajo un modelo asíncrono centrado en la Plataforma Databricks. El flujo técnico es el siguiente:

- **Capa de Presentación solicitud de reporte (Interfaz de Usuario):** Los usuarios interactúan con una Databricks App, (*una aplicación web desarrollada con frameworks compatibles como Streamlit, Dash, Gradio, entre otros*) que muestra pantallas con los reportes definidos, permite iniciar una solicitud de alguno de los reportes definidos con parámetros que diligencia el usuario final y permite hacer seguimiento al estado de ese reporte solicitado.
- **Procesamiento por Lotes (ETL):** Pipelines Databricks, desarrollados en notebooks con PySpark y/o SQL, se ejecutan de forma procesada (jobs). Estos procesos leen las solicitudes de forma programada (en lotes) usando la(s) Delta Table(s) disponibles en los medallones plata u oro *(y en caso de no existir en alguno de estos medallones como alternativa final medallón bronce)*, consultan los datos requeridos de las capas de la arquitectura medallón y aplican las transformaciones y agregaciones necesarias.
- **Almacenamiento en Capa Gold:** Los resultados de los pipelines, ya consolidados y enriquecidos, se ***materializan*** como nuevas Delta Tables en la capa Gold de la arquitectura medallón. Estas tablas están optimizadas para consultas analíticas de alto rendimiento y servirán para generar archivos csv y html.

- **Consumo de Resultados:** Los usuarios acceden a los resultados a través de tres mecanismos principales, todos gobernados por Unity Catalog:
    - **Databricks App:** Permite consultar el **estado de las solicitudes registradas** (por ejemplo: pendiente, en proceso, finalizada),
    - **Espacio de trabajo Databricks** (menú catálogo, para usuarios sin conocimientos técnicos): descargar archivos resultado guardados en volúmenes. (Html y csv)
    - **Espacio de trabajo Databricks** (usuarios de negocio avanzados con conocimientos mínimos en sql / pyspark): mediante notebooks pueden explorar los conjuntos de datos de reportes csv o consultar las tablas en etapa gold.

- **Componentes Tecnológicos Clave**
    - **Plataforma de Cómputo y Orquestación**: Azure Databricks es el núcleo de la solución, proveyendo los clústeres de cómputo para el procesamiento distribuido, el orquestador de jobs para los pipelines y el entorno de ejecución para la aplicación y los dashboards.
    - **Almacenamiento de Datos:** Se utilizará Azure Storage (StorageV2 - uso general v2) como el sistema de archivos subyacente para el Lakehouse. Todos los datos en las capas Medallón (Bronze, Silver, Gold) se almacenarán en formato Delta Lake, garantizando transacciones ACID, versionado de datos y rendimiento.
    - **Gobernanza y Seguridad de Datos**: Unity Catalog es el componente central para la gobernanza. Se utilizará para:
        - **Control de Acceso Basado en Roles (RBAC):** Definir grupos de usuarios y asignar permisos granulares sobre catálogos, esquemas, tablas y vistas.
        - **Autenticación y Autorización**: El acceso a la solución estará federado a través del servicio de identidad de la entidad (Microsoft Entra ID), gestionado mediante Unity Catalog.
        - **Linaje de Datos y Auditoría**: Aprovechar las capacidades nativas de Unity Catalog para trazar el flujo de datos y auditar los accesos.

6.2. Principios de Diseño Técnico

- **Tecnología Nativa**: La solución se construye exclusivamente con herramientas y servicios disponibles en la Plataforma Databricks y Azure, evitando dependencias externas que requieran licenciamiento adicional o complejas integraciones. Los resultados se materializan como archivos csv y html que pueden ser consultados en cualquier sistema operativo por parte del usuario final solicitante.
- **Inmutabilidad y Atomicidad:** El uso de Delta Lake asegura la integridad de los datos en todas las capas. Los procesos ETL generarán tablas en la capa Gold, nunca modificarán los datos de origen en las capas Bronze o Silver.
- **Desacoplamiento:** La arquitectura asincrónica desacopla la interfaz de usuario de los procesos de cómputo intensivo, garantizando que la experiencia del usuario sea fluida y que los recursos de procesamiento se utilicen de manera eficiente.
- **Reusabilidad y Mantenibilidad:** Los pipelines de datos y las plantillas de notebooks se diseñarán de forma modular para ser reutilizables. La gestión del código se realizará mediante repositorios de control de versiones usando GIT a través de azure devops, permitiendo un ciclo de vida de desarrollo estructurado CI/CD.
- **Sistema <u>No Transaccional</u>:** Se reitera que la solución es un sistema analítico de consulta y no un sistema transaccional (OLTP). **<u>No interactúa en tiempo real</u>** con los sistemas de origen de documentos electrónicos ***ni modifica datos fuera de su propio***

flujo en el Lakehouse. (Son procesos por lotes que tendrán periodicidad de 1 hora dentro de los horarios laborales).

- **Sistema No Transaccional:** Se reitera que la solución es un sistema de ***registro de solicitudes con procesamiento por lotes*** y no un sistema transaccional (OLTP***). No interactúa en tiempo real*** con los sistemas de origen de documentos electrónicos ni modifica datos fuera de su propio flujo en el Lakehouse. (Los procesos por lotes tendrán periodicidad de 1 hora dentro de los horarios laborales).

7. Fases ejecución

Esta sección NO representa un cronograma ni plan de trabajo, el equipo contratista tendrá que realizar los análisis necesarios, sesiones internas de su equipo y en general, las labores que considere pertinentes para definir tales elementos.

Esta sección presenta de forma general las fases y entregables para poder definir el proceso comercial.

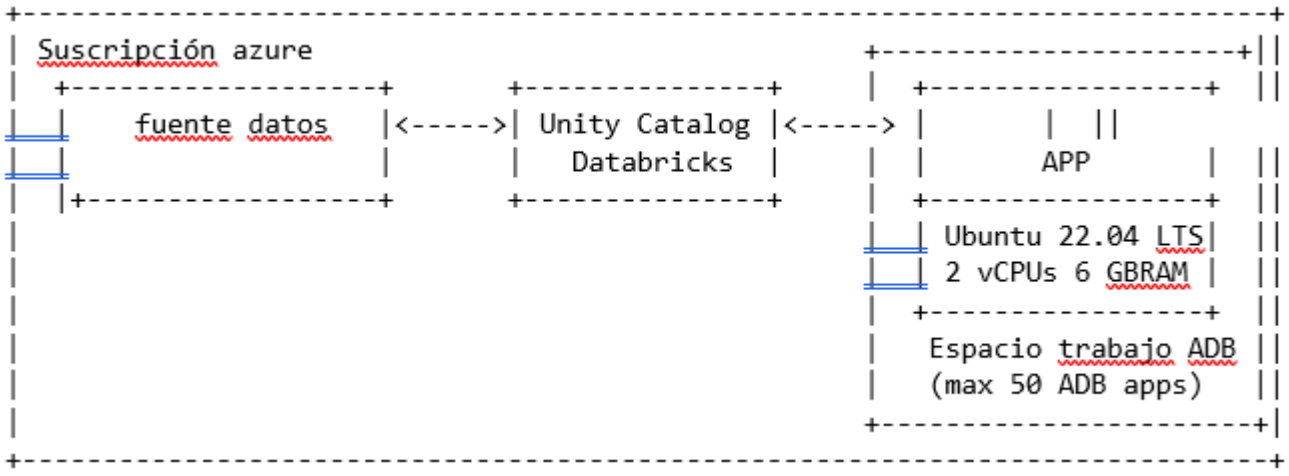
No.	Fase	Descripción de Actividades Principales	Entregables
1	**Inicio y Diseño**	• Reunión (o reuniones) de inicio, validación del plan de trabajo y configuración de entornos. • Sesiones intensivas con SFESO para priorizar y definir las especificaciones detalladas de los reportes clave (PMV). • Diseño de los modelos de datos Gold (definiciones y características de las tablas) para el PMV y maquetas de la App/Dashboards.	• Plan de trabajo detallado • Especificaciones funcionales del PMV • Diseño técnico y maquetas
2	**Desarrollo Iterativo**	• Desarrollo de los pipelines de datos para el PMV en ciclos cortos (iteraciones por definir entre 1 y 3 semanas). • Construcción del esqueleto de la Databricks App. • Construcción del esqueleto de las ETL • Construcción del esqueleto de plantillas html para reportes materializados a través de las ETL • Alineación con la seguridad a nivel de fila (Row-Level Security). • Pruebas unitarias y de integración continuas por parte del contratista. • Validaciones de calidad según especificaciones por áreas de negocio.	• Entrega de la App con funcionalidad de solicitud y visualización de estado gestión solicitud. • Entrega de ETLs mediante notebooks. • Entrega de plantillas html para visualización de resultados que permitan a las ETL generar los reportes html reales.
3	**Pruebas, Estabilización y Despliegue**	• Ejecución de pruebas de rendimiento y seguridad sobre el PMV. • **Pruebas de Aceptación de Usuario (UAT)** realizadas por la SFESO y usuarios finales de otras áreas cuando el reporte no contenga exclusivamente datos originados en los conjuntos de documentos electrónicos.	• Informe de resultados de pruebas incluyendo las pruebas unitarias y de integración, de sistema, end -to- end, de carga/stress, funcionales, validaciones de calidad de datos y cualquier tipo

		• Corrección de defectos identificados en UAT. • Despliegue supervisado de la solución MVP en el entorno productivo.	de prueba realizada, finalizando con la aceptación de usuario (UAT) • Solución PMV con todos sus componentes (interfaz para registro de solicitudes como app databricks, ETL's) desplegada en producción.
4	**Capacitación y Cierre**	• Ejecución del plan de capacitación funcional y técnico. • Entrega de toda la documentación final del proyecto (manuales, guías). • Firma del acta de entrega y cierre del proyecto (FT-IIT-2708).	• Sesiones de transferencia de conocimiento, profundización, y demás actividades necesarias con sus respectivos Certificados y/o actas de capacitación. • Paquete documental final entregado. • Acta de Cierre de Proyecto.

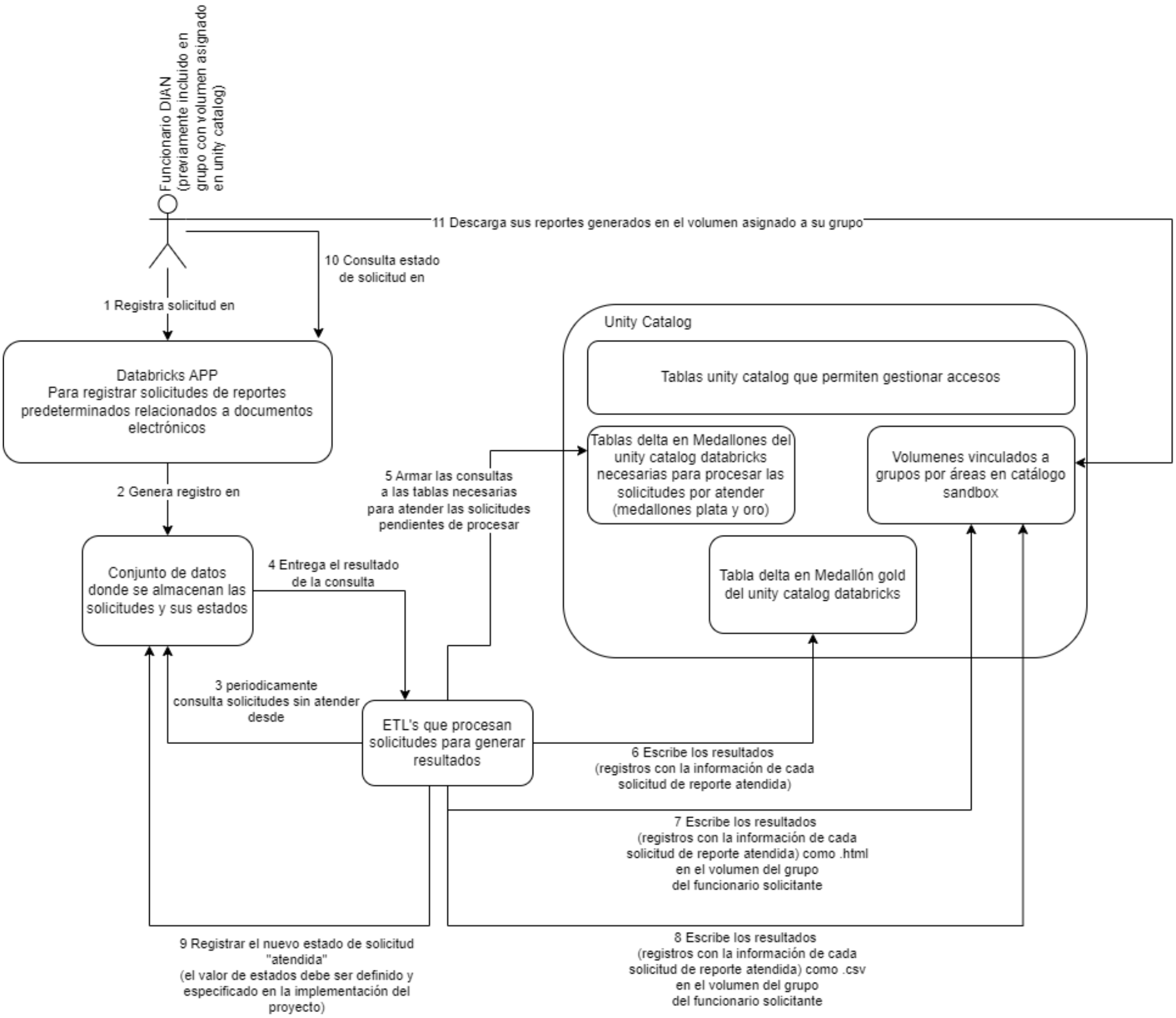
8. Diagramas de proceso de negocio

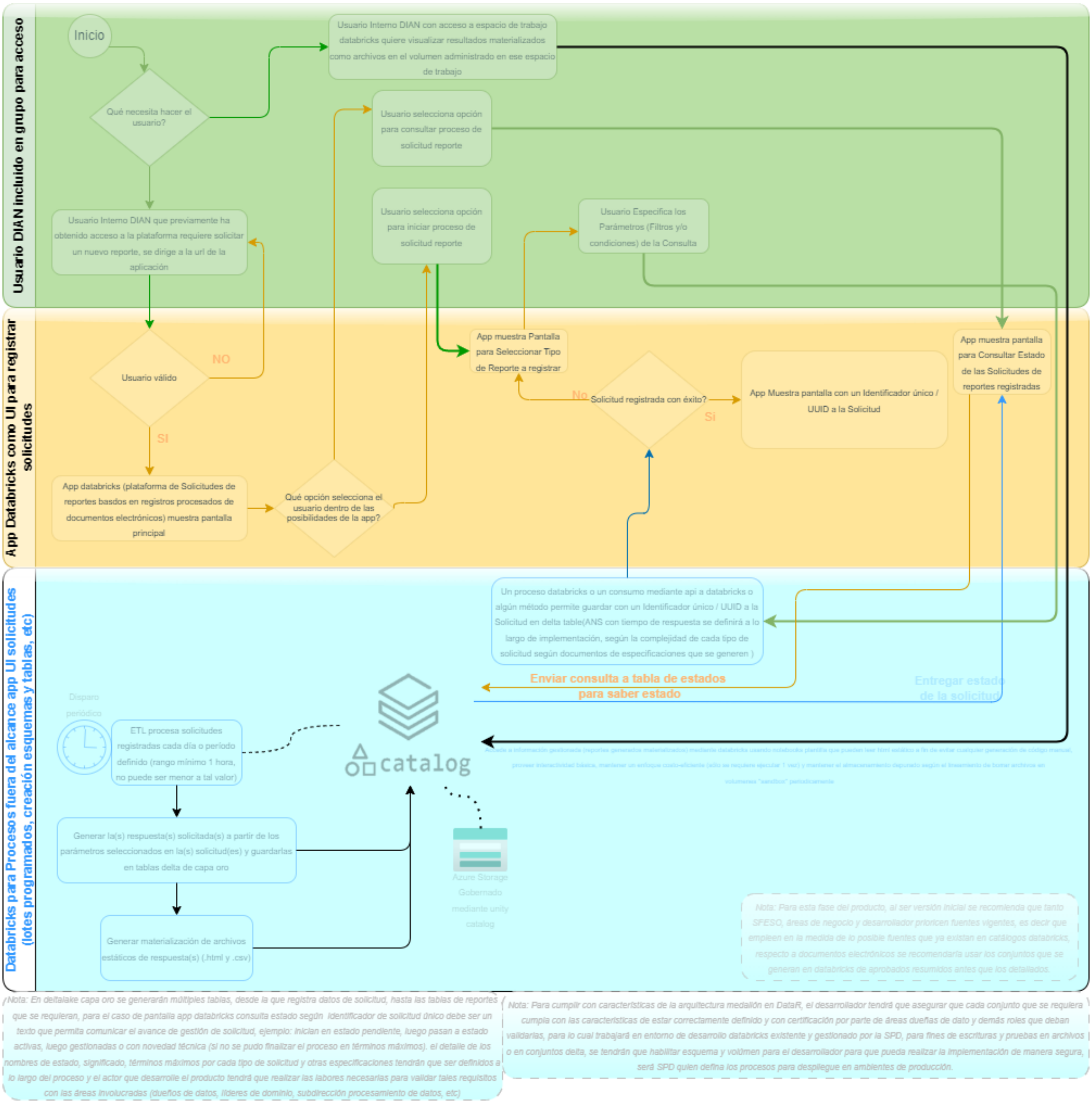
8.1. Diagramas de componentes databricks app (enfoque "Data entry forms backed by Databricks SQL")

8.1.1. Diagramas de proceso Databricks APP



8.1.2. Diagrama general proceso para registrar solicitud de reporte y consultar estado solicitud




Usuario DIAN incluido en grupo para acceso
Inicio
Qué necesita hacer el usuario?
Usuario Interno DIAN con acceso a espacio de trabajo databricks quiere visualizar resultados materializados como archivos en el volumen administrado en ese espacio de trabajo
Usuario selecciona opción para consultar proceso de solicitud reporte
Usuario selecciona opción para iniciar proceso de solicitud reporte
Usuario Especifica los Parámetros (Filtros y/o condiciones) de la Consulta
Usuario Interno DIAN que previamente ha obtenido acceso a la plataforma requiere solicitar un nuevo reporte, se dirige a la url de la aplicación
App Databricks como UI para registrar solicitudes
Usuario válido
NO
SI
App muestra Pantalla para Seleccionar Tipo de Reporte a registrar
No
Solicitud registrada con éxito?
Si
App Muestra pantalla con un Identificador único / UUID a la Solicitud
App muestra pantalla para Consultar Estado de las Solicitudes de reportes registradas
App databricks (plataforma de Solicitudes de reportes basdos en registros procesados de documentos electrónicos) muestra pantalla principal
Qué opción selecciona el usuario dentro de las posibilidades de la app?
Un proceso databricks o un consumo mediante api a databricks o algún método permite guardar con un Identificador único / UUID a la Solicitud en delta table(ANS con tiempo de respuesta se definirá a lo largo de implementación, según la complejidad de cada tipo de solicitud según documentos de especificaciones que se generen)
Enviar consulta a tabla de estados para saber estado
Entregar estado de la solicitud
Databricks para Procesos fuera del alcance app UI solicitudes (lotes programados, creación esquemas y tablas, etc)
Disparo periódico
ETL procesa solicitudes registradas cada día o período definido (rango mínimo 1 hora, no puede ser menor a tal valor)
catalog
Generar la(s) respuesta(s) solicitada(s) a partir de los parámetros seleccionados en la(s) solicitud(es) y guardarlas en tablas delta de capa oro
Generar materialización de archivos estáticos de respuesta(s) (.html y .csv)
Azure Storage Gobernado mediante unity catalog
Nota: Para esta fase del producto, al ser versión inicial se recomienda que tanto SFESO, áreas de negocio y desarrollador prioricen fuentes vigentes, es decir que empleen en la medida de lo posible fuentes que ya existan en catálogos databricks, respecto a documentos electrónicos se recomendaría usar los conjuntos que se generan en databricks de aprobados resumidos antes que los detallados.
Nota: En deltalake capa oro se generarán múltiples tablas, desde la que registra datos de solicitud, hasta las tablas de reportes que se requieran, para el caso de pantalla app databricks consulta estado según identificador de solicitud único debe ser un texto que permita comunicar el avance de gestión de solicitud, ejemplo: inician en estado pendiente, luego pasan a estado activas, luego gestionadas o con novedad técnica (si no se pudo finalizar el proceso en términos máximos). el detalle de los nombres de estado, significado, términos máximos por cada tipo de solicitud y otras especificaciones tendrán que ser definidos a lo largo del proceso y el actor que desarrolle el producto tendrá que realizar las labores necesarias para validar tales requisitos con las áreas involucradas (dueños de datos, líderes de dominio, subdirección procesamiento de datos, etc)
Nota: Para cumplir con características de la arquitectura medallón en DataR, el desarrollador tendrá que asegurar que cada conjunto que se requiera cumpla con las características de estar correctamente definido y con certificación por parte de áreas dueñas de dato y demás roles que deban validarlas, para lo cual trabajará en entorno de desarrollo databricks existente y gestionado por la SPD, para fines de escrituras y pruebas en archivos o en conjuntos delta, se tendrán que habilitar esquema y volúmen para el desarrollador para que pueda realizar la implementación de manera segura, será SPD quien defina los procesos para despliegue en ambientes de producción.

9. Reportes

9.1. Reporte inicial que debe ser definido para identificar si es posible incluir en versión inicial de la solución o requiere generación de conjuntos Gold adicionales en declaraciones relacionadas

Deben generarse las especificaciones detalladas del reporte, su implementación dependerá de que se cumplan como pre-requisitos:

- Existencia preliminar de fuentes nube pertinentes, cada una con las condiciones necesarias para realizar las consultas u operaciones definidas
- Validación de calidad de tales conjuntos por parte de los dueños de datos en cada caso (validaciones funcionales que realizaría SFESO junto con el equipo contratista)

1	**Reporte comparativo valores Declaraciones de IVA vs valores en documentos electrónicos**
La interfaz de registro de solicitudes debe permitir seleccionar un tipo de informe para comparar valores de la declaración de IVA (Formulario 300) de un contribuyente, ingresando el NIT y el periodo. La vista debe mostrar un resumen de los conceptos declarados y permitir la descarga del formulario completo para su análisis detallado.	

Para el reporte, hay que definir qué elementos debe contener cada pantalla para registro de solicitud, los filtros a aplicar, la estructura del resultado final, identificación específica de cada dato (origen, tipo de dato, etc) necesario para generarlo, relaciones, operaciones que apliquen, y en general el detalle completo individual. **Su implementación dependerá de la disponibilidad de tiempo y de conjuntos origen previamente existentes y validados (certificados por área de negocio)**, en caso de no ser posible su implementación en el tiempo de este proceso, los entregables deberán ser la documentación completa, modelos de datos, (ETL's) parametrizables en origenes y destinos para entorno de desarrollo, aplicando estrategia de fuentes sintéticas tipo objetos simulados (mocking) con sus respectivas pruebas unitarias y plantilla de reporte html con datos sintéticos compatibles con los objetos simulados en la ETL correspondiente para posteriormente implementar en otro proceso de contratación posterior.

Cualquier reporte tendrá como máximo 5000 registros (filas), desde la interfaz de la databricks app y en los archivos html debe existir un texto al respecto.

9.2. Reportes complementarios que deben ser definidos para probar la solución, cumpliendo estándares internos (aplicando formatos de especificación y de registro pruebas)

El detalle (especificación para cada informe / reporte) se daría en sus respectivos formatos documentales (Word historias de usuario y Excel especificación de reporte) que elaboraría el proveedor en conjunto con un usuario de área de negocio cliente y SFESO.

La implementación del proceso para generar cada informe / reporte de forma automatizada en lotes será analizada en mesas de trabajo periódicas (semanales) para definir si el proveedor implementa su ETL en notebooks dentro de un espacio de trabajo Databricks con acceso a volumen Sandbox que luego deben ser validados por SPD e integrados en las ETL y conjuntos que crea o si la implementa directamente el equipo SPD según la disponibilidad de recursos.

1	**Reporte a nivel detallado de los 200 registros top superior e inferior en facturación electrónica para un NIT receptor en un periodo específico**
La interfaz de registro de solicitudes debe permitir seleccionar un tipo de informe para la generación asíncrona de un reporte detallado del top de facturación electrónica (100 registros de mayor valor	

facturado y 100 registros de menor valor facturado), para lo cual al solicitarse se dispondrán como filtros: el NIT del emisor, el NIT del receptor y un rango de fechas en un periodo limitado empleando para ello los filtros de fechas que garanticen consultas repetibles/confiables (proceso determinista). Los resultados deben visualizarse en una tabla. La tabla debe tener una cantidad de registros que pueda visualizarse en celda de notebook databricks completa (que no requiera paginación). La tabla debe mostrar campos clave como tipo de documento, CUFE, número de factura, divisa, forma y medio de pago, fechas de emisión y firma, NIT y razón social del emisor y adquiriente, ciudad de venta, impuestos discriminados (IVA, ICA, INC, etc.), retenciones, subtotales, descuentos y valor total de la factura, entre otros.	
2	**Reporte perfil global de operaciones de ingresos del Contribuyente (EMISOR)**
La interfaz de registro de solicitudes debe permitir seleccionar un tipo de informe para analizar el comportamiento del emisor en términos de ingresos. Esta debe incluir datos consolidados como la cantidad y el valor total de facturas generadas en un periodo limitado empleando para ello los filtros de fechas que garanticen consultas repetibles/confiables (proceso determinista). Además, se deben generar listados tipo “TOP 10” de compradores según valor y cantidad de facturas, así como los principales conceptos de retenciones e impuestos facturados. También se debe incluir un listado de notas débito y crédito generadas, con posibilidad de ver las notas asociadas a cada factura.	
3	**Reporte perfil global de operaciones de egresos del Contribuyente (ADQUIRIENTE)**
La interfaz de registro de solicitudes debe permitir seleccionar un tipo de informe orientado a analizar las compras realizadas por un contribuyente en un periodo limitado empleando para ello los filtros de fechas que garanticen consultas repetibles/confiables (proceso determinista). Esta debe incluir tarjetas con la cantidad y el valor de facturas recibidas, y listados “TOP 10” de proveedores, IVA pagado, retenciones aplicadas e impuestos pagados. La información debe estar organizada de forma que facilite la identificación de patrones de comportamiento y posibles riesgos fiscales.	
4	**Reporte top relaciones entre contribuyentes**
La interfaz de registro de solicitudes debe permitir seleccionar un tipo de informe que permita consultar el top 5000 de las relaciones comerciales entre dos contribuyentes, mediante la búsqueda por NIT del emisor y del adquiriente. Esta consulta debe mostrar cuántos documentos se han generado (total de documentos emitidos) y cuántos se han recibido (cantidad de documentos a nombre del adquiriente), permitiendo identificar relaciones recurrentes o significativas que puedan ser relevantes para la fiscalización.	

Para estos elementos además de la documentación completa debe entregarse funcionando en la app databricks en entorno desarrollo las respectivas pantallas.

Para cada uno de estos reportes, la especificación detallada deberá incluir: las métricas y cálculos exactos, los filtros requeridos (ej. búsqueda de uno o varios NITs), las visualizaciones (gráficos, tablas), y las reglas para la seguridad a nivel de fila (ej. por seccional o área de negocio).

Cualquier tipo de reporte tendrá como máximo 5000 registros (filas).
Cualquier tipo de reporte está sujeto a los límites de procesamiento para los procesos ETL (disparos cada 60 minutos, con los recursos disponibles en databricks)

10. Requerimientos funcionales y no funcionales

10.1. RQF-01 Requerimientos funcionales Databricks APP para registro de solicitudes de reportes predeterminados federado por unity catalog

Los usuarios serán funcionarios de la DIAN válidos en unity catalog vinculados a un grupo con volumen asignado. La APP debe identificar las características del grupo para identificar su volumen y permisos de acceso a conjuntos. Según los permisos de acceso podrá visualizar las pantallas para solicitud de reportes compatibles.

En los RQF 01.1 a RAF 01.7 se especifican las generalidades de la databricks app que permitirá a los usuarios interesados en solicitar reportes predeterminados registrar sus solicitudes a fin de que sean atendidas automáticamente por lotes programados.

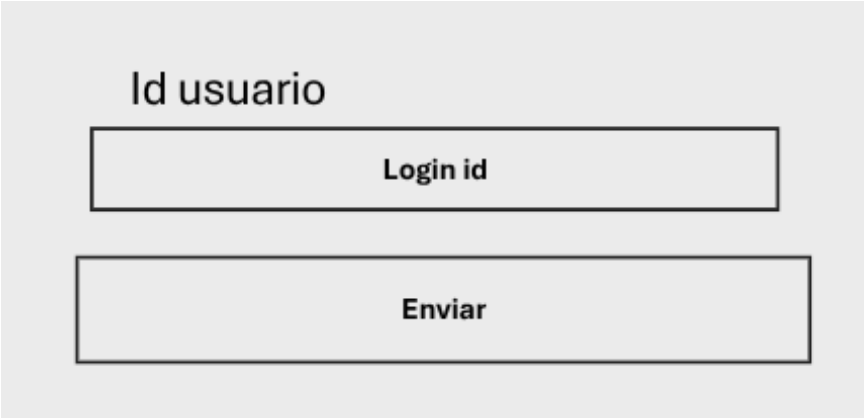
10.1.1. RQF-01.1 Ingreso Databricks APP para registro de solicitudes de reportes predeterminados federado por unity catalog

El ingreso a la app debe estar condicionado por unity catalog, permitiendo acceso exclusivamente a usuarios que estén incluidos en los grupos que se hayan creado para tal finalidad.

- ***Precondiciones:***
  - Usuario registrado en un grupo en unity catalog.
  - Servicio para validar identidad está disponible y operativo con normalidad al momento de iniciar ingreso.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - Se muestran las opciones pertinentes para seleccionar uno de los posibles informes o reportes.
  - Al seleccionar un informe o reporte, el comportamiento del sistema es el esperado (se muestra en la pantalla la caracterización de ese tipo de informe)

- ***Wireframe ilustrativo:***



10.1.2. RQF-01.2 Menú de navegación Databricks APP para registro de solicitudes de reportes predeterminados

Mediante opciones de menú debe ser posible seleccionar una de las posibles características o módulos disponibles, como lo son:

- Registrar solicitud

- Consultar estado de solicitudes
- Ayuda

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - Se muestran las opciones pertinentes para seleccionar uno de los posibles módulos
  - Al seleccionar un informe o reporte, el comportamiento del sistema es el esperado (se muestra en la pantalla el módulo seleccionado).

- ***Wireframe ilustrativo:***



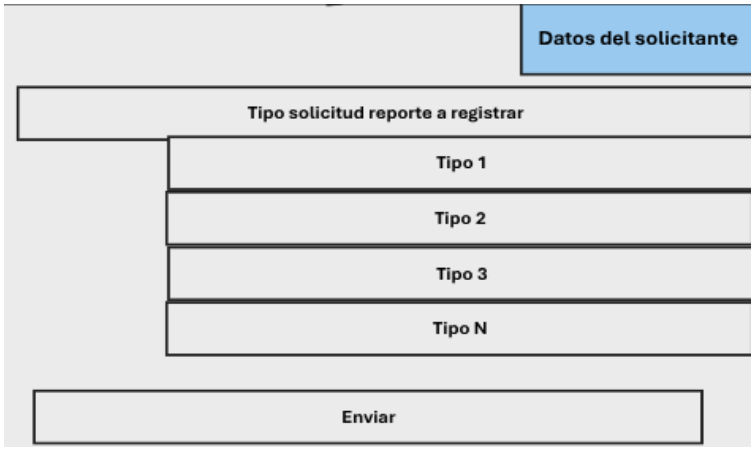
10.1.3. RQF-01.3 Selección de tipos de informes/reportes disponibles Databricks en la APP para registro de solicitudes de reportes predeterminados

Mediante opciones de menú debe ser posible seleccionar uno de los posibles informes o reportes disponibles, el elemento seleccionado permite registrar las características de la solicitud a registrar.

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - Se muestran las opciones pertinentes para seleccionar uno de los posibles informes o reportes.
  - Al seleccionar un informe o reporte, el comportamiento del sistema es el esperado (se muestra en la pantalla la caracterización de ese tipo de informe)

- ***Wireframe ilustrativo:***



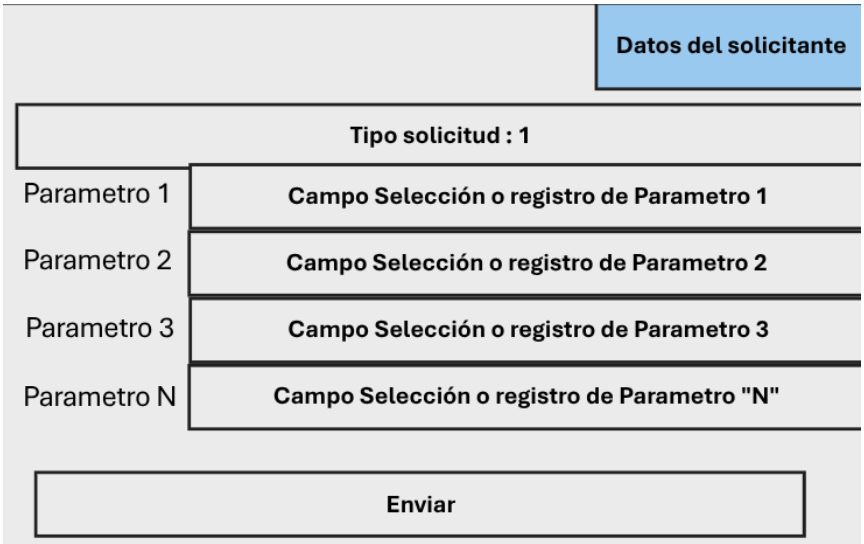
10.1.4. RQF-01.4 Captura de parámetros de la solicitud en la Databricks APP para registro de solicitudes de reportes predeterminados

Mediante campos con opciones seleccionables o para registrar valores (texto, número o fecha) se caracteriza la solicitud del tipo de reporte deseado.

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - Se muestran las opciones pertinentes para caracterizar uno de los posibles informes o reportes. (Tales opciones son específicas de ese tipo de informe o reporte)
  - Al seleccionar un informe o reporte, el comportamiento del sistema es el esperado (se muestra en la pantalla la caracterización de ese tipo de informe)

- ***Wireframe ilustrativo:***



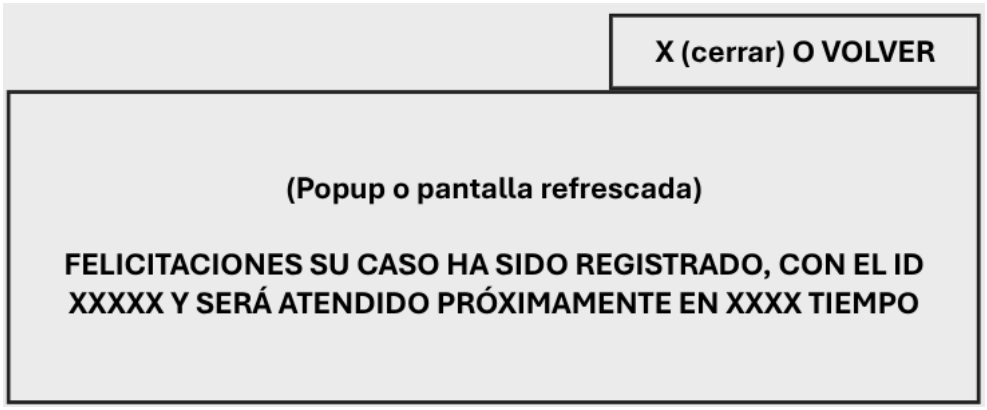
10.1.5. RQF-01.5 Confirmación de estado registro de solicitud en la Databricks APP para registro de solicitudes de reportes predeterminados

Mediante elementos en la pantalla debe ser posible dar retroalimentación de los posibles estados de registro de solicitud (exitosa o fallida).

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - El estado de registro de solicitud debe ser consistente con su registro real en la fuente de datos donde se almacenan las solicitudes (si el registro se creó, debe aparecer en la fuente de datos).
  - En caso de registro exitoso debe presentar el identificador de la solicitud registrada para poder hacer seguimiento.
  - En caso de fallo debe presentar un código para poder reportar el error.
  - Debe existir opción de cerrar / regresar a la interfaz principal.

- ***Wireframe ilustrativo:***



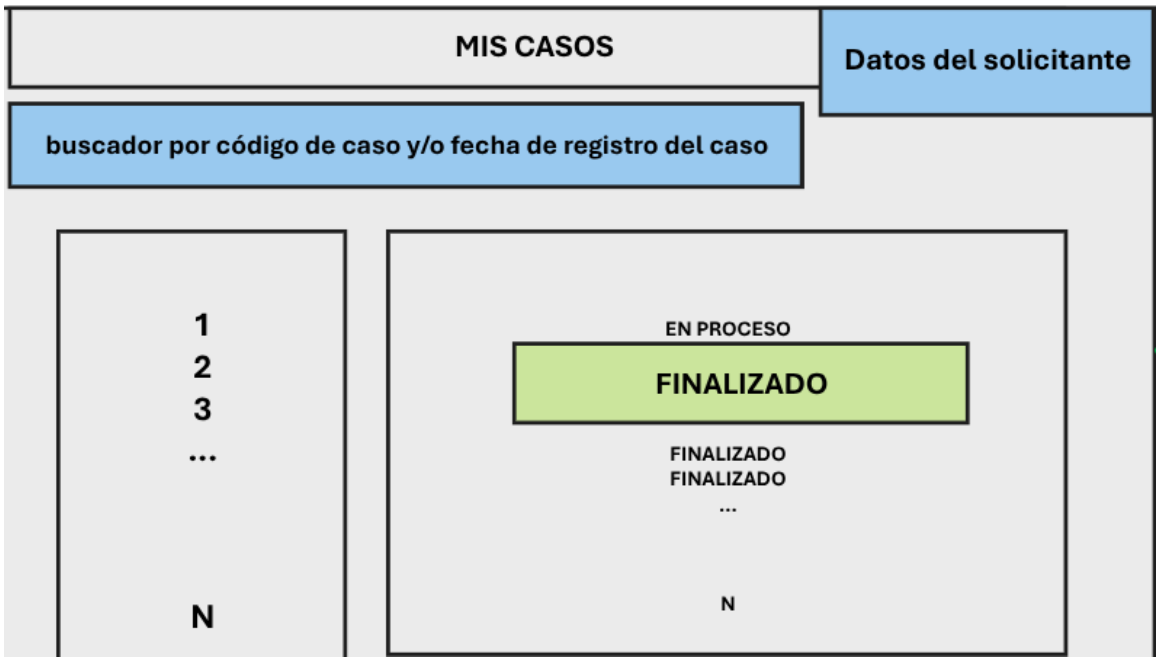
10.1.6. RQF-01.6 Consulta de estado solicitudes registradas

Mediante elementos en la pantalla debe ser posible conocer el estado de los posibles estados de avance de atención a solicitudes registradas.

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - El estado de registro de solicitud debe ser consistente con su registro real en la fuente de datos donde se almacenan las solicitudes (si el registro se creó, debe aparecer en la fuente de datos).
  - Debe ser posible filtrar por identificador de caso o fecha de registro de caso

- ***Wireframe ilustrativo:***



10.1.7. RQF-01.7 Sección de ayuda, preguntas frecuentes y guías de uso para usuarios en la Databricks APP para registro de solicitudes de reportes predeterminados

Mediante elementos en la pantalla debe ser posible brindar herramientas orientadas a la experiencia de usuario que permitan al solicitante conocer cómo usar la herramienta, sus capacidades, limitaciones, preguntas frecuentes, etc.

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet.

- ***Criterios de aceptación:***
  - El usuario válido que ha ingresado puede acceder a la pantalla y visualizar los contenidos en los distintas pantallas o contenidos que se definan con el experto UX y área de negocio.
  - En caso de fallo debe presentar un código para poder reportar el error.
  - Debe existir opción de cerrar / regresar a la interfaz principal.

10.2. RQF-02 Requerimientos reportes materializados.

Las solicitudes registradas mediante databricks APP almacenadas en una fuente que el contratista determine según su análisis técnico contemplando eficiencia, tiempo de respuesta, disponibilidad, etc, que podrá ser una base de datos relacional (postgresql), una tabla delta, con las características que determine... será la fuente de origen para procesar tales solicitudes que generarán archivos como materialización de los contenidos solicitados disponible para descarga mediante unity catalog, en los volumenes de cada grupo según los lineamientos internos de gobernanza existentes.

A continuación se detallan los RQF que definen las generalidades de los resultados materializados como archivos html y csv y su origen en tabla delta dentro de la capa gold.

10.2.1. RQF-02.1 Materialización de resultados en tabla delta

Los resultados de los reportes solicitados a través de la app databricks deben ser almacenados como nuevos registros en tablas delta dentro de la capa gold, que servirán tanto para replicar los resultados de una solicitud a través de cuadernos mediante consultas sql y para generar los archivos que podrá descargar el solicitante a través del volumen de su grupo en unity catalog.

- ***Precondiciones:***
  - Acceso confirmado al catálogo de databricks previamente solicitado y gestionado.
  - Conectividad de red a internet (para acceder al espacio de trabajo databricks en donde se usará el cuaderno).
  - Se ha definido la estrategia técnica y el contenido del reporte a atender de manera que se asegure no superará las 5000 filas (registros)
  - Existen los elementos de consulta original necesarios para poder generar los resultados solicitados.

- ***Criterios de aceptación:***
  - Los resultados de las solicitudes se almacenan correctamente en las tablas gold definidas para tal fin.
  - El resultado se puede consultar de forma eficiente (cada solicitud atendida entregará los resultados mediante consulta sql en tiempos no superiores a 2 minutos incluso si se usa un cluster de propósito general de capacidad mínima -single node-.)
  - A medida que se escriban resultados en las tablas delta se mantiene el rendimiento esperado, es decir, no se presenta degradación del conjunto a medida que crece pues el contratista ha implementado un modelo de datos enfocado en consultas eficientes.

10.2.2. RQF-02.2 Acceso a resultados materializados como archivos html y .csv mediante volumen en unity catalog (usuarios finales sin conocimientos técnicos)

Mediante la interfaz de databricks se podrán descargar los resultados materializados (archivo .html y archivo .csv) que podrá usar el solicitante para visualizar sin necesidad de tener conocimientos técnicos (archivo .html) o para importar como archivo plano en excel (para el caso del .csv).

Pantalla unity catalog, recurso esquema habilitado para el grupo de trabajo en donde existe un volumen a donde se escriben los archivos.

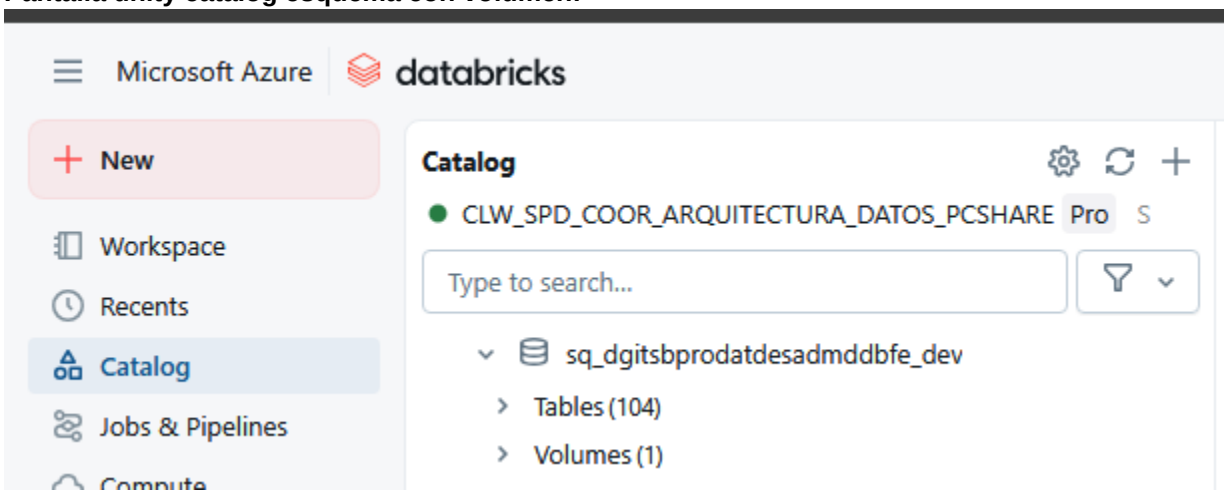
- ***Precondiciones:***
  - Acceso confirmado al catálogo de databricks previamente solicitado y gestionado.
  - Acceso a volumen de área dentro de databricks previamente solicitado y gestionado.
  - Conectividad de red a internet (para acceder al espacio de trabajo databricks en donde se usará el cuaderno).
  - Contenido del reporte en cualquiera de sus variantes (html y csv) no puede ser superior a 10MB.
  - Los vínculos externos que puedan requerirse (imágenes, scripts javascript, etc) deben ser llamados mediante https y estar almacenados dentro de los recursos DIAN o en portales que hayan sido catalogados como seguros (por ejemplo para una libreria como d3.js validar si se considera un repositorio seguro al momento de implementar).

- ***Criterios de aceptación:***

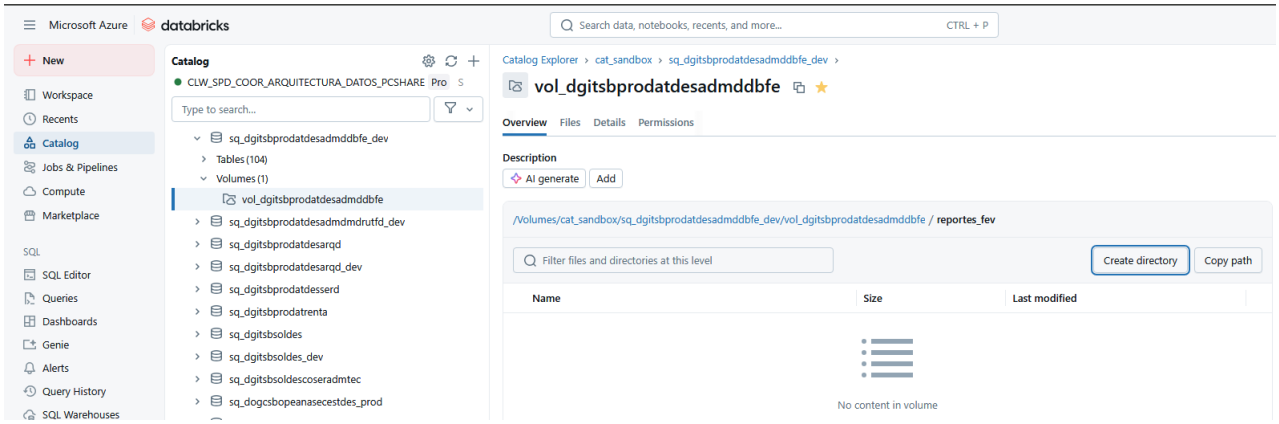
- El resultado html se puede visualizar en los navegadores disponibles en equipos de cómputo de funcionarios DIAN en instalaciones físicas DIAN.
- El resultado csv se puede importar a excel sin generar inconvenientes

Pantallas ilustrativas

Pantalla unity catalog esquema con volumen:

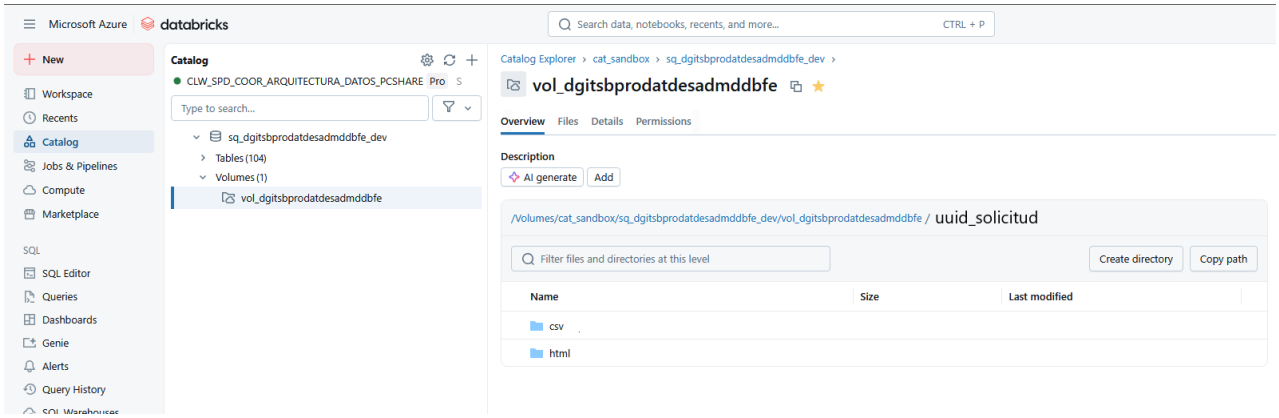


Pantalla unity catalog exploración contenidos en el volumen donde se podrá descargar el informe materializado:

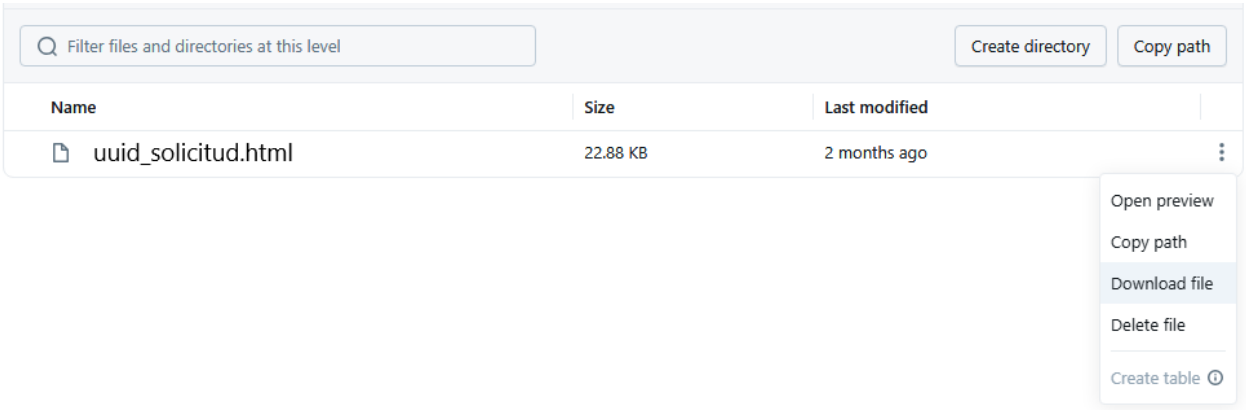


Pantalla consulta contenidos volumen:

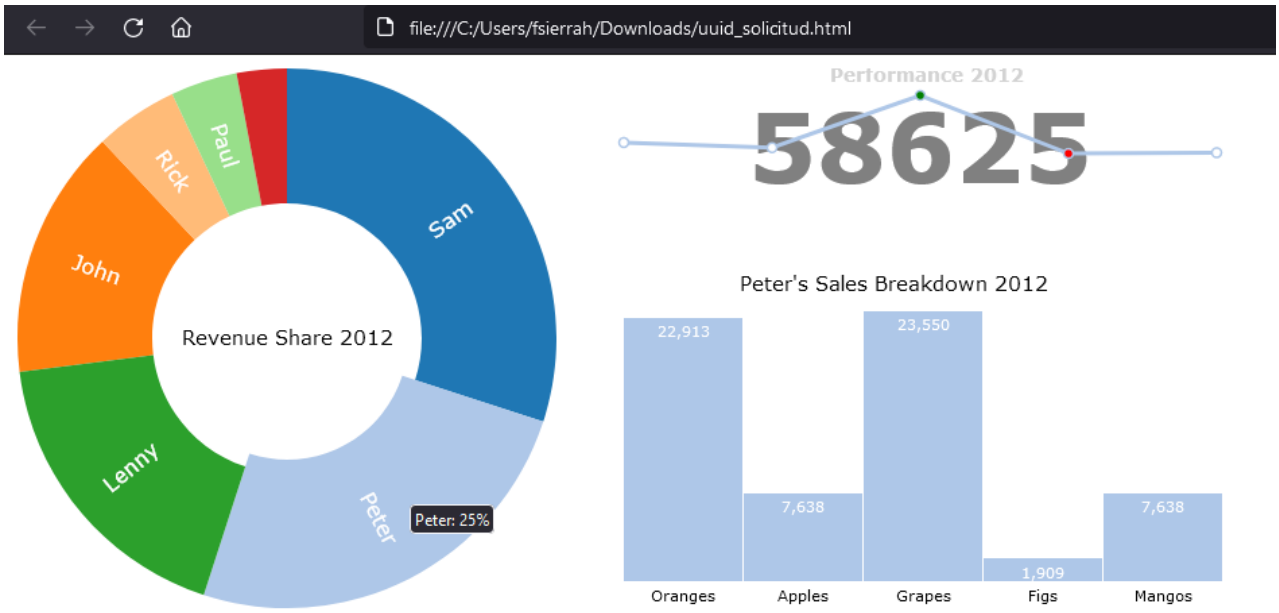
Cuando se hayan generado reportes, podrán descargarse directamente desde el volumen:



Pantalla Descarga del archivo html desde el volumen en unity catalog:

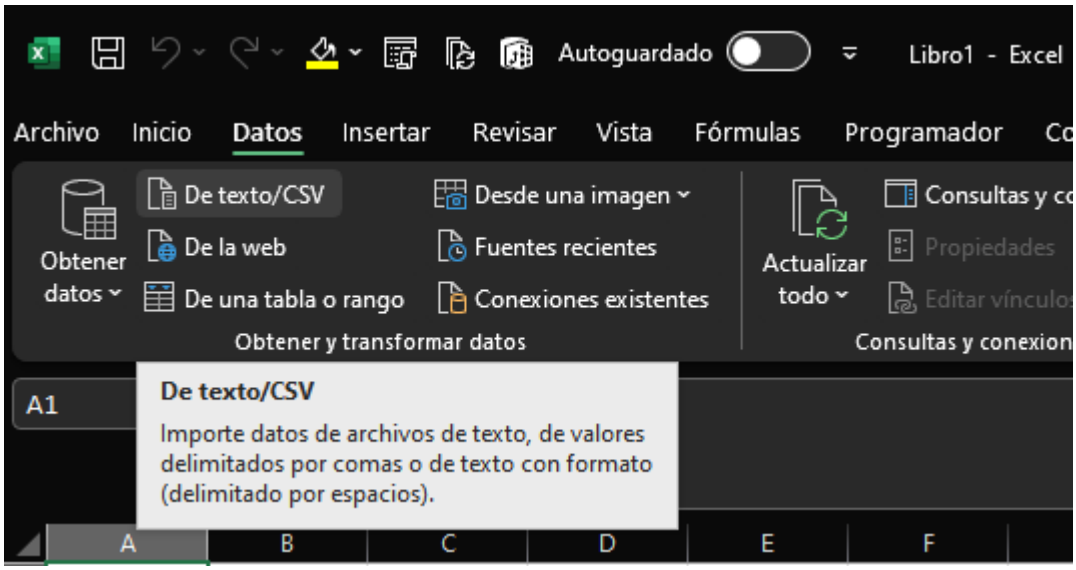


Pantalla Visualización del html descargado en un navegador del usuario final:

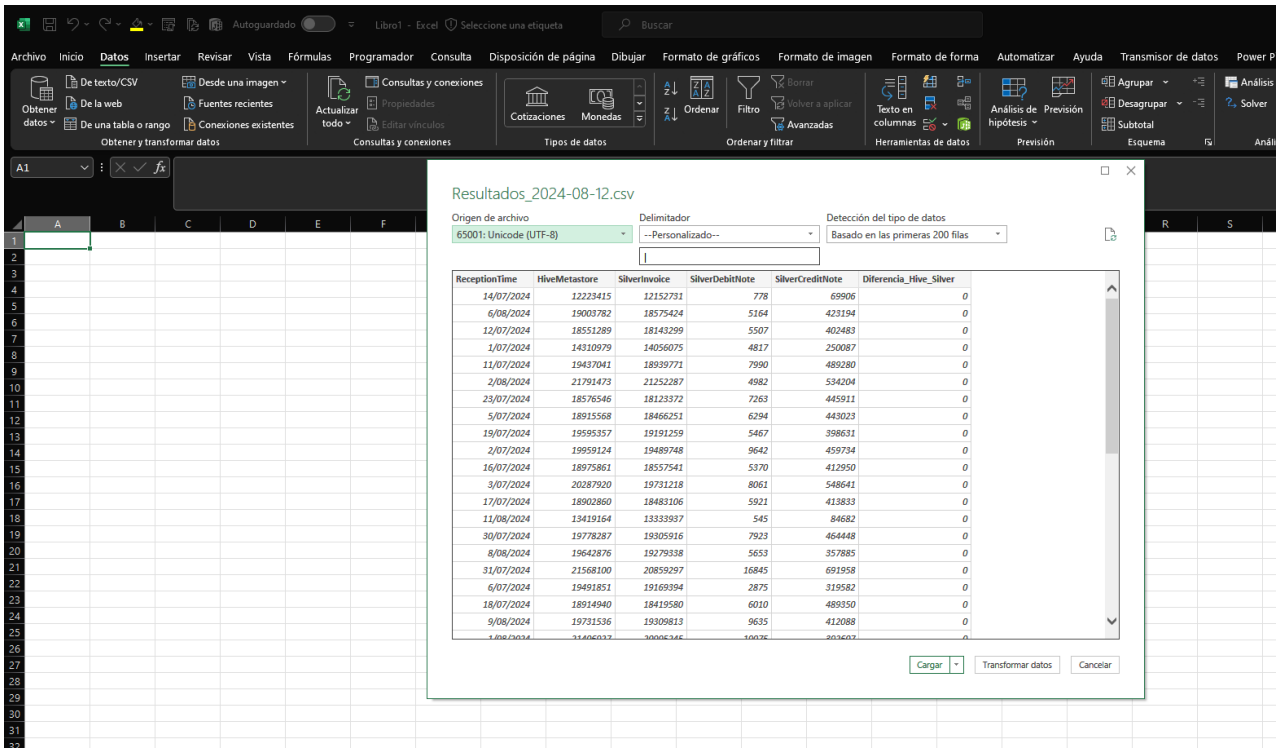


Cuando el usuario final desea trabajar en excel con los resultados csv, una vez descargue el archivo deberá importarlo y podrá trabajar con las herramientas propias de ese software (modelos, tablas dinámicas, etc)

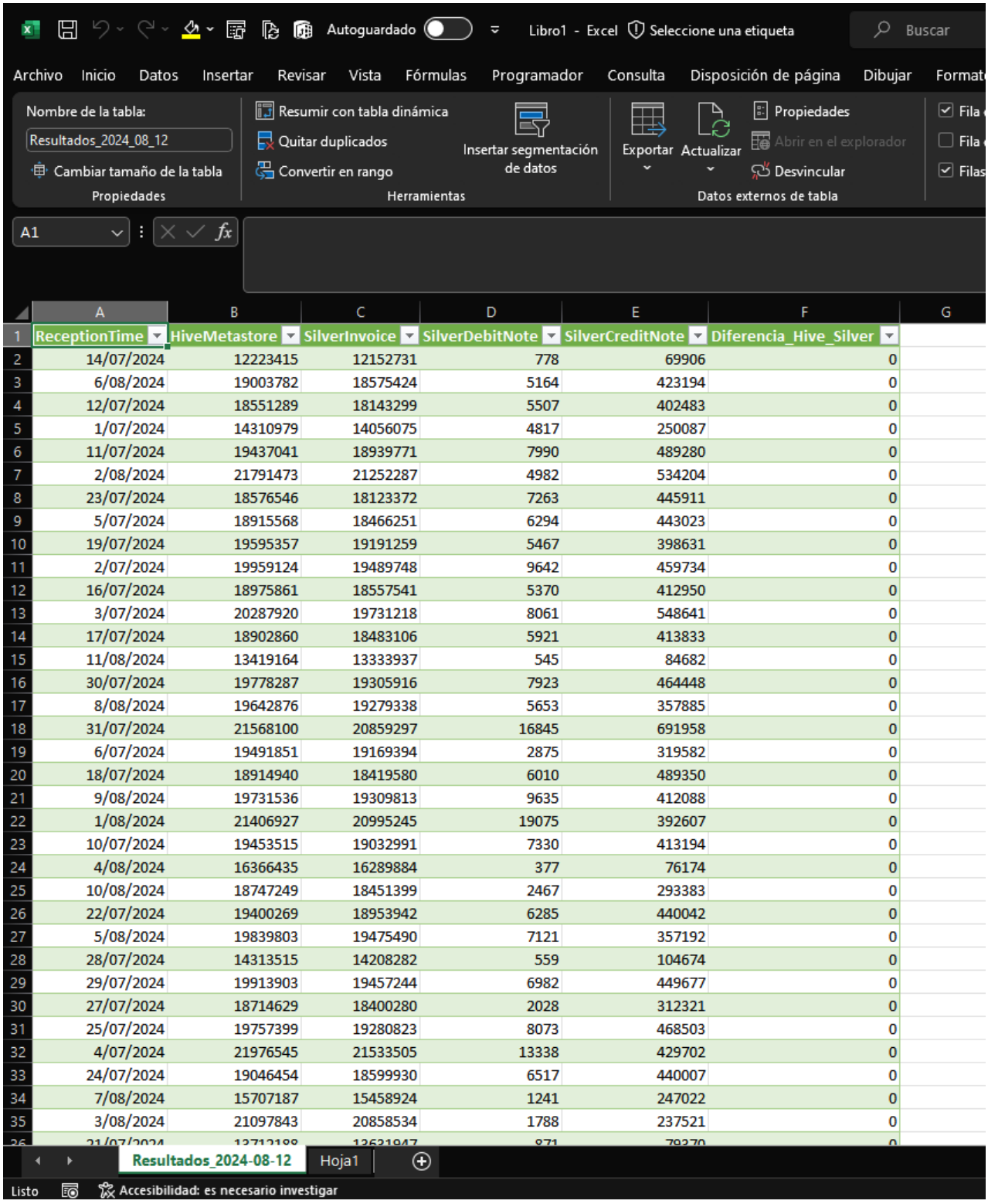
Pantalla uso de menú para importar archivo csv descargado:



Diálogo excel para seleccionar codificación, delimitador, etc antes de proceder con la carga en el libro:



Pantalla ejemplo libro poblado con contenidos del csv descargado (tabla excel generada):



Autoguardado Libro1 - Excel Seleccione una etiqueta Buscar

Archivo Inicio Datos Insertar Revisar Vista Fórmulas Programador Consulta Disposición de página Dibujar Formato

Nombre de la tabla: Resultados_2024_08_12
Cambiar tamaño de la tabla
Propiedades
Resumir con tabla dinámica
Quitar duplicados
Convertir en rango
Herramientas
Insertar segmentación de datos
Exportar
Actualizar
Propiedades
Abrir en el explorador
Desvincular
Datos externos de tabla

A1

	A	B	C	D	E	F	G
1	ReceptionTime	HiveMetastore	SilverInvoice	SilverDebitNote	SilverCreditNote	Diferencia_Hive_Silver	
2	14/07/2024	12223415	12152731	778	69906	0	
3	6/08/2024	19003782	18575424	5164	423194	0	
4	12/07/2024	18551289	18143299	5507	402483	0	
5	1/07/2024	14310979	14056075	4817	250087	0	
6	11/07/2024	19437041	18939771	7990	489280	0	
7	2/08/2024	21791473	21252287	4982	534204	0	
8	23/07/2024	18576546	18123372	7263	445911	0	
9	5/07/2024	18915568	18466251	6294	443023	0	
10	19/07/2024	19595357	19191259	5467	398631	0	
11	2/07/2024	19959124	19489748	9642	459734	0	
12	16/07/2024	18975861	18557541	5370	412950	0	
13	3/07/2024	20287920	19731218	8061	548641	0	
14	17/07/2024	18902860	18483106	5921	413833	0	
15	11/08/2024	13419164	13333937	545	84682	0	
16	30/07/2024	19778287	19305916	7923	464448	0	
17	8/08/2024	19642876	19279338	5653	357885	0	
18	31/07/2024	21568100	20859297	16845	691958	0	
19	6/07/2024	19491851	19169394	2875	319582	0	
20	18/07/2024	18914940	18419580	6010	489350	0	
21	9/08/2024	19731536	19309813	9635	412088	0	
22	1/08/2024	21406927	20995245	19075	392607	0	
23	10/07/2024	19453515	19032991	7330	413194	0	
24	4/08/2024	16366435	16289884	377	76174	0	
25	10/08/2024	18747249	18451399	2467	293383	0	
26	22/07/2024	19400269	18953942	6285	440042	0	
27	5/08/2024	19839803	19475490	7121	357192	0	
28	28/07/2024	14313515	14208282	559	104674	0	
29	29/07/2024	19913903	19457244	6982	449677	0	
30	27/07/2024	18714629	18400280	2028	312321	0	
31	25/07/2024	19757399	19280823	8073	468503	0	
32	4/07/2024	21976545	21533505	13338	429702	0	
33	24/07/2024	19046454	18599930	6517	440007	0	
34	7/08/2024	15707187	15458924	1241	247022	0	
35	3/08/2024	21097843	20858534	1788	237521	0	

Resultados_2024-08-12 Hoja1

Listo Accesibilidad: es necesario investigar

10.2.3. RQF-02.3 Acceso a resultados mediante notebooks

Para los usuarios avanzados será posible que mediante notebooks databricks (cuadernos jupyter) se consulten los resultados de los distintos tipos de solicitud usando como origen los archivos csv del volumen, o las tablas del conjunto gold si el usuario hace parte de un grupo con acceso de lectura.

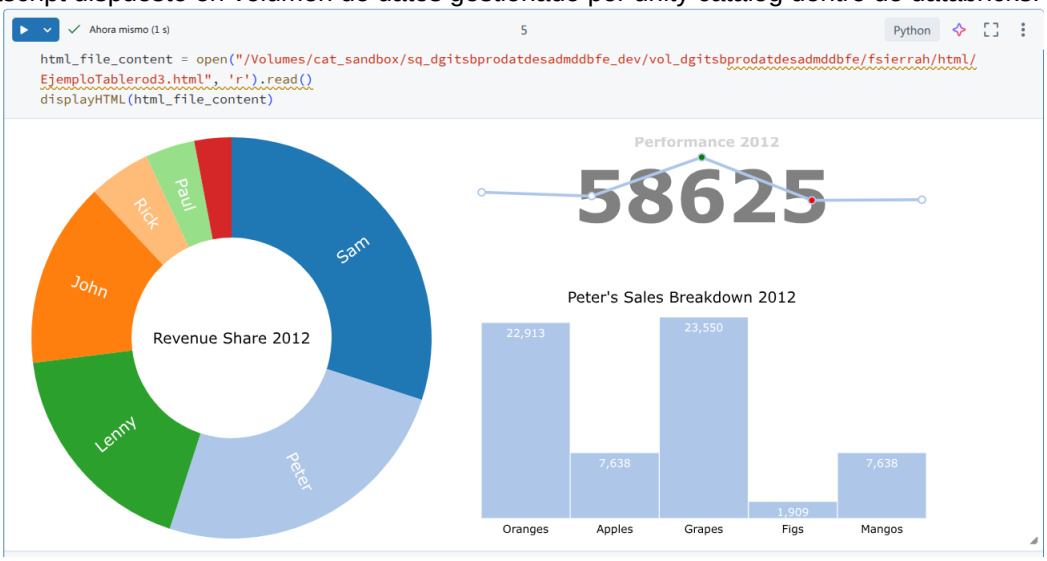
Podrán existir plantillas de ejemplo que permitan a usuarios avanzados basarse en tales elementos que contengan el código necesario para acceder a los datos disponibles en catálogo databricks, nivel medallón oro relacionados a la solicitud atendida.

- ***Precondiciones:***
  - Usuario válido ha ingresado al servicio.
  - Conectividad de red a internet (para acceder al espacio de trabajo databricks en donde se usará el cuaderno).
  - Acceso al catálogo de databricks previamente solicitado y gestionado.
  - Acceso a volumen de área dentro de databricks previamente solicitado y gestionado.
  - El cuaderno como visualizador de resultados o reporte materializado se realiza basado en python para la ejecución de celdas, e internamente puede mostar html empleando para interactividad Javascript, SVG, ya sea como código directamente registrado en celdas o como archivo externo ubicado en el volumen de área.
  - Contenido del reporte no puede ser superior a 10MB.
  - Los vínculos externos que puedan requerirse (imágenes, scripts javascript, etc) deben ser llamados mediante https y estar almacenados dentro de los recursos DIAN o en portales que hayan sido catalogados como seguros (por ejemplo para una libreria como d3.js validar si se considera un repositorio seguro al momento de implementar).

- ***Criterios de aceptación mínimos:***
  - El cuaderno renderiza los resultados esperados en html en una celda aplicando la instrucción displayHTML(htmlCode) satisfactoriamente (sin errores).
  - El cuaderno que renderiza resultados se limita a mostrar lo registrado en la solicitud de forma estática, presentando tablas, elementos visuales como gráficos o textos en diversos tamaños de manera que cumplan lo definido en los documentos de especificaciones que se realicen para el tipo de solicitud específico.

- ***Wireframe ilustrativo consulta archivo html generado:***

Pantalla de ejemplo visualización en notebook databricks que renderiza archivo .html con interactividad por javascript dispuesto en volumen de datos gestionado por unity catalog dentro de databricks:



Para los csv o consultas a tablas en "gold" tendrá que leerse el contenido como un dataframe y posteriormente el usuario podrá operar como guste con tal estructura.

10.3. Requerimientos no funcionales de los ETLs que procesarán los datos de origen, realizarán las transformaciones o agregaciones necesarias y guardarán los resultados de reportes atendidos tanto en tablas delta como en archivos html y csv.

10.3.1. RNF-01 Seguridad

Definición	Acceso exclusivo a funcionarios permitidos definidos dentro de Unity Catalog
Requerimiento funcional impactado	RQF-01.6 Acceso a resultados materializados como archivos html y .csv mediante volumen en unity catalog (usuarios finales sin conocimientos técnicos)
Excepción	Ninguna
Método prueba	Acceso a la app restringido vinculado al nombre de usuario en dominio (Entra ID), a través de los mecanismos que SPD indique al proveedor están disponibles dentro de Databricks (por ejemplo: a través de registro mediante app registration en azure y key vaults, uso del "Unified Data Orquestator", etc) *Se tendrá que validar usuarios que no deben tener acceso efectivamente no pueden ver o ejecutar los flujos (workflows / jobs) que se implementen*
Unidad de Medida	Cumple o No Cumple
Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple

10.3.2. RNF-02 Disponibilidad

Definición	Sólo se disparan procesos ETL originados en la app exclusivamente 7:00 a 19:00 zona horaria Bogotá, Colombia
Requerimiento funcional impactado	RQF-01.6 Acceso a resultados materializados como archivos html y .csv mediante volumen en unity catalog (usuarios finales sin conocimientos técnicos)
Excepción	Fallo de infraestructura nube (si los servicios están offline) Fallo de conexión de red del usuario final (cliente)
Método prueba	App no disponible fuera de horario indicado
Unidad de medida	Cumple/No Cumple
Criterio de aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple.

10.4. Requerimientos no funcionales de la app Databricks como interfaz para registro de solicitudes de reportes asíncronas

10.4.1. RNF-01 Seguridad

Definición	Acceso exclusivo a funcionarios permitidos definidos dentro de Unity Catalog
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Ninguna
Método prueba	Acceso a la app restringido vinculado al nombre de usuario en dominio (Entra ID), a través de los mecanismos que SPD indique al proveedor están disponibles dentro de Databricks (por ejemplo: a través de registro mediante app registration en azure y key vaults, uso del “Unified Data Orquestator”, etc) *Se tendrá que validar usuarios que no deben tener acceso efectivamente no pueden ver la opción de app, ni ingresar a la app con URL específica.*
Unidad de Medida	Cumple o No Cumple
Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple

10.4.2. RNF-02 Disponibilidad

Definición	Acceso a la app exclusivamente en horario laboral (7:00 a 18:00 zona horaria Bogotá, Colombia)
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Fallo de infraestructura nube (si los servicios están offline) Fallo de conexión de red del usuario final (cliente)
Método prueba	App no disponible fuera de horario indicado
Unidad de medida	Cumple/No Cumple
Criterio de aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple.

10.4.3. RNF-03 Usabilidad: interfaz consistente y con comportamiento predecible

Definición	Pantallas siguen el flujo definido, con comportamiento esperado y existen elementos visuales de ayuda para el usuario final según se hayan definido en las historias de usuario necesarias. Se hace uso de los recursos oficiales relacionados a interfaz de usuario en la entidad y los elementos como logos o fuentes son parametrizables en caso de que requieran actualización. El sistema debe manejar adecuadamente los errores y excepciones, proporcionando mensajes claros y comprensibles al usuario en caso de fallos durante la ejecución.
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Ninguna
Método prueba	Al navegar, los usuarios experimentan un flujo consistente (el comportamiento entre pantallas y al usar elementos disponibles es repetible). No hay iconos, logos ni elementos de entidades distintas a la DIAN
Unidad de Medida	Cumple o No Cumple

Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple

10.4.4. RNF-04 Desempeño: tiempo de respuesta

Definición	Cada opción que se presenta en pantalla tiene un tiempo de respuesta menor a 2 minutos
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento (mostrar pantallas, menús, guardar una solicitud, mostrar el estado de solicitudes registradas)
Excepción	Fallo de infraestructura nube (si los servicios están offline) Fallo de conexión de red del usuario final (cliente)
Método prueba	Cada pantalla responde dentro del tiempo definido
Unidad de medida	Cumple/No Cumple
Criterio de aceptación	El porcentaje de cumplimiento deberá ser como mínimo del 95%. Las pruebas de aceptación deberán arrojar un SI cumple.

10.4.5. RNF-05 Desempeño: consumo de recursos

Definición	La App debe funcionar consumiento recuros inferiores a las capacidades máximas de las app databricks (menor a 6GB memoria) y para registrar solicitudes no requieren iniciar clusters adicionales.
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Fallo de infraestructura nube (si los servicios están offline) Fallo de conexión de red del usuario final (cliente)
Método prueba	En escenario de concurrencia máxima de 200 usuarios simultáneos, la herramienta permite operar con la capacidad de recursos disponibles sin presentar fallos.
Unidad de medida	Cumple/No Cumple
Criterio de aceptación	El porcentaje de cumplimiento deberá ser como mínimo del 95%. Las pruebas de aceptación deberán arrojar un SI cumple.

10.4.6. RNF-06 Restricciones de diseño: tecnologías empleadas

Definición	App databricks (debe usar las capacidades y características disponibles para apps databricks), no se permite el uso de componentes adicionales comerciales (librerías, frameworks que requieran licencias por uso ni pagos adicionales)
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Ninguna
Método prueba	El producto está implementado en un espacio de trabajo databricks que provea SPD Se pueden emplear como fuentes para salvado de registros “transaccionales” las que se definan, que pueden ser tablas delta, azure table, postgresql integrado en databricks Una vez existan conjuntos, deben accederse desde databricks y mantener las reglas definidas en unity catalog
Unidad de Medida	Cumple o No Cumple

Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple. *Nota: Si se usa una fuente con enfoque relacional en cualquier parte del proceso, se recomienda el uso de OLTP Database en databricks con postgresql* 

10.4.7. RNF-07 Rendimiento: usuarios concurrentes

Definición	Debe garantizarse que al menos 200 usuarios pueden tener acceso a la app de forma simultánea sin que genere fallo, indisponibilidad o afecten el rendimiento del producto.
Requerimiento funcional impactado	Aplica para todas las funcionalidades de app. Todos los especificados en este documento
Excepción	Fallo de infraestructura nube (si los servicios están offline) Fallo de conexión de red del usuario final (cliente)
Método prueba	App no disponible fuera de horario indicado App funcional para la cantidad de usuarios concurrentes definidos y según lo indicado en RQF Desempeño: Consumo de recursos.
Unidad de medida	Cumple/No Cumple
Criterio de aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un SI cumple.

10.5. Requerimientos no funcionales para visualización por HTML:

10.5.1. RNF-10 Visualización HTML: compatibilidad con navegadores modernos

Definición	Las visualizaciones generadas en HTML deben ser compatibles con los navegadores modernos utilizados por la entidad (Microsoft Edge, Google Chrome, Mozilla Firefox). El contenido debe renderizarse correctamente sin errores de formato ni pérdida de funcionalidad.
Requerimiento funcional impactado	RQF-02.1, RQF-02.2
Excepción	Ninguna
Método prueba	Visualización del contenido HTML en los navegadores definidos por la entidad, validando que los elementos interactivos y visuales se muestran correctamente.
Unidad de Medida	Cumple / No Cumple
Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un “Sí cumple”.

10.5.2. RNF-11 Visualización HTML: uso de recursos seguros y auditables

Definición	Todo recurso externo referenciado en el HTML (scripts, hojas de estilo, imágenes) dentro de los volúmenes de datos gestionados por

	Unity Catalog. No se permite el uso de recursos de terceros no auditables.
Requerimiento funcional impactado	RQF-02.1, RQF-02.2
Excepción	Ninguna
Método prueba	Revisión del código fuente HTML para verificar que todos los recursos externos cumplen con las políticas de seguridad.
Unidad de Medida	Cumple / No Cumple
Criterio de Aceptación	El porcentaje de cumplimiento deberá ser del 100%. Las pruebas de aceptación deberán arrojar un “Sí cumple”.

10.5.3. RNF-12 Visualización HTML: rendimiento y tiempo de carga

Definición	Las visualizaciones HTML deben cargarse completamente en el navegador web en un tiempo máximo de 5 segundos para archivos de hasta 10MB. El contenido debe mostrarse sin bloqueos ni retardos perceptibles para el usuario.
Requerimiento funcional impactado	RQF-02
Excepción	Fallos de red o infraestructura fuera del control de la app.
Método prueba	Medición del tiempo de carga desde la ejecución de la celda o apertura de la vista hasta la visualización completa del contenido.
Unidad de Medida	Cumple / No Cumple
Criterio de Aceptación	El porcentaje de cumplimiento deberá ser como mínimo del 95%. Las pruebas de aceptación deberán arrojar un “Sí cumple”.

11. Reglas de negocio

ID	Regla de Negocio
RN01	Solo los usuarios autorizados, pertenecientes a grupos definidos en Unity Catalog en Databricks, pueden acceder a los recursos disponibilizados y consultar información consolidada de Documentos Electrónicos de acuerdo con la solicitud registrada.
RN02	Las consultas deben ejecutarse de forma asincrónica, permitiendo que el usuario no tenga que esperar en tiempo real por los resultados pues pueden o no estar disponibles en conjuntos existentes, tampoco genera picos de consumo innecesarios pues los recursos son empleados de forma eficiente mediante procesos por lotes y en general, ETLs que atienden los casos a gestionar de forma independiente al funcionamiento de la app databricks para registrar solicitudes de reportes.
RN03	Se recomienda que los procesos de ETL sean viables para ejecutar cada 60 minutos.
RN04	Cuando se empleen notebooks para visualizar resultados, será posible permitir la descarga de los resultados de las consultas en formatos CSV, siempre que exista previamente un volumen de datos para el área solicitante que así lo permita y en la especificación detallada de ese tipo de reporte se haya indicado necesidad de generar el csv en el volumen.

RN05	Las solicitudes de consulta deben estar basadas en parámetros definidos previamente por el usuario que garanticen un proceso determinista. Ej. NIT del emisor, NIT del receptor, tipo de documento electrónico y rango de fechas.
RN06	El sistema debe mostrar mensajes de estado durante la ejecución de las consultas, indicando si están en progreso, finalizadas o si ocurrió un error.
RN07	En caso de errores críticos durante la ejecución de un proceso lote para atención de consultas, se soluciona internamente en SPD sin que afecte la experiencia de usuario final, si puede poner en riesgo términos definidos para entrega de algún conjunto, el estado de la solicitud deberá poder servir como indicador de novedad técnica para el usuario solicitante.
RN08	1. La solución debe garantizar la integridad y trazabilidad de los datos consultados, sin permitir su modificación desde la interfaz del usuario.
RN09	Los resultados de las consultas deben visualizarse en al menos uno de los siguientes canales: 1. Descargables desde interfaz databricks unity catalog (contenidos de volumen). 2. Notebook Databricks.
RN10	La app debe ser reutilizable y mantenible, permitiendo su actualización sin afectar la experiencia del usuario final.
RN11	La solución no debe generar ni modificar documentos electrónicos; su función es exclusivamente de consulta y análisis de información ya registrada.
RN12	La plataforma de solicitud de reportes debe permitir registrar el tipo de reporte requerido y los parámetros de consulta, sirviendo como punto de entrada para el proceso.
RN13	La solución debe cumplir con las políticas de seguridad de la entidad en cuanto al manejo de datos sensibles y confidenciales.
RN14	La solución no permitirá solicitar informes / reportes que no hayan sido previamente definidos y validados por áreas interesadas (solicitantes y SFESO) y disponibilizados por SPD.
RN15	Los reportes de resultados se generarán con un máximo de 5000 registros.

12. Estrategia de Calidad, Pruebas y Criterios de Aceptación

Los requerimientos funcionales y no funcionales presentados pueden ser complementados mediante las especificaciones de reportes individuales, historias de usuario y en general, documentación de se genere a lo largo del proceso a fin de dar mayor detalle a cada elemento.

La aceptación formal de la solución por parte de la DIAN se regirá por una estrategia de pruebas por niveles, culminando con la validación de que todos los requerimientos funcionales y no funcionales definidos han sido cumplidos con la aprobación de las áreas funcionales y/o técnicas según corresponda.

12.1. Niveles QA (aseguramiento de calidad)

El contratista es responsable de *ejecutar y documentar* todos los niveles de QA antes de presentar la solución para la validación final por parte de la DIAN:

Calidad de datos:

- Validaciones de completitud en los conjuntos simulados.

- Evaluación de calidad técnica (cantidad de nulos, repetidos, etc).
- Todas las labores relacionadas al aseguramiento de calidad incluyendo consistencia de datos y calidad de los conjuntos según las definiciones y reglas de negocio.

Pruebas Internas del Contratista:

Como mínimo se esperan las siguientes pruebas:

- **Pruebas Unitarias y de Integración:** Verificación de que cada componente de software (databricks app) y pipelines de datos (etl's) funciona correctamente de forma individual y en conjunto.
- **Pruebas del Sistema (End-to-End):** El contratista debe ejecutar pruebas completas que simulen los flujos de usuario, desde la solicitud del reporte hasta su visualización, para verificar internamente el cumplimiento de los Requerimientos Funcionales.
- **Pruebas de carga/stress:** El contratista deberá simular uso simultáneo de la app y ejecutar flujos de atención de solicitudes simuladas (sin intervención humana).

Pruebas de Aceptación de Usuario (UAT):

- **Responsabilidad:** Este hito de validación final será liderado y ejecutado por SFESO, y un representante de un área que consumirá ese tipo de reporte (ejemplo: un funcionario de recaudo que requiere el informe con datos del tipo de reporte #1 participa como usuario final para las pruebas funcionales corerspondientes).
- **Alcance de las UAT:** SFESO validará un conjunto de casos de uso de negocio sobre un entorno de pruebas estable, provisto por el contratista con datos de prueba pertinentes y previamente acordados. (Ejemplo: registran solicitudes, y ven en la app que se muestra el estado de avance, cuando finalice procesamiento luego de 1 hora podrán visualizar los resultados generados tanto como archivo html como csv en el volumen dispuesto para su grupo, también podrán consultar en las tablas delta de capa gold los resultados de su solicitud específica y no podrán consultar resultados de otras solicitudes de otras áreas, al comprobar los resultados procesados son consistentes en todos los medios y contienen la información esperada, el resultado html dispuesto en el volumen se visualiza según lo especificado.)

12.2. Criterios de Aceptación

La solución será formalmente aceptada (culminando con la firma del Acta de Entrega, FT-IIT-2708) únicamente cuando se demuestre el cumplimiento de los siguientes criterios. Cada criterio se validará contra los requerimientos específicos ya definidos en este documento:

- **Descripción:** SFESO validará que la solución implementa todas las funcionalidades descritas en la Sección 9 (Reportes) y sección 10 (Requerimientos Funcionales).
  - **Método de Verificación:** Durante las pruebas UAT, SFESO validará que cada uno de los requerimientos operan como se espera. El contratista deberá presentar una matriz de trazabilidad que vincule cada RQF con su respectivo caso de prueba.
- **Criterio de Aceptación de los Reportes (MVP):**
  - **Descripción**: SFESO validará para dar aceptación sobre los reportes definidos como entregables iniciales:

- Implementados según especificaciones,
- Los datos presentados deben ser correctos y consistentes en todas sus formas de generación (tabla delta, csv y html).

- **Método de Verificación**: SFESO (que puede solicitar apoyo de la SIA) validará la precisión de los datos y cálculos de cada reporte del MVP. La consistencia de los datos y la correcta aplicación de las reglas de negocio definidas son un factor de aceptación no negociable, se registrará la aceptación o rechazo en documento firmado por las partes.

- **Criterio de Aceptación No Funcional:**
  - **Descripción**: La solución debe cumplir con todos los atributos de calidad, rendimiento y seguridad definidos en la Sección 11 (Requerimientos No Funcionales).
  - **Método de Verificación**: El contratista deberá entregar informes de prueba que evidencien el cumplimiento de, como mínimo, los siguientes RNF clave:
  - **Seguridad (RNF-01):** Pruebas que demuestren el correcto funcionamiento del acceso restringido por Unity Catalog.
  - **Disponibilidad (RNF-02):** Verificación de que la aplicación respeta el horario laboral definido.
  - **Rendimiento y Concurrencia (RNF-04, RNF-07):** Reporte de pruebas de carga que demuestre la estabilidad de la Plataforma con 150 usuarios simultáneos y tiempos de respuesta dentro de los umbrales definidos.
  - **Usabilidad (RNF-03):** La validación de la usabilidad se realizará de forma cualitativa durante las pruebas UAT por parte de SFESO.

- **Criterio de Aceptación de Documentación y Capacitación:**
  - **Descripción**: Toda la documentación y las sesiones de capacitación definidas en la Sección 13 (Plan de Capacitación) deben haber sido entregadas y aprobadas por la DIAN.
  - **Método de Verificación**: La DIAN emitirá una comunicación formal de aprobación de los manuales y actas de asistencia a las capacitaciones.

- **Criterio de Cierre Formal:**
  - **Descripción**: El hito final de aceptación es la firma del acta de pruebas UAT.
  - **Método de Verificación**: SFESO deberá firmar el acta de conformidad de las Pruebas de Aceptación de Usuario (formato FT-IIT-2712). La existencia de defectos menores no bloqueantes puede ser aceptada, siempre que se acuerde un plan para su resolución por escrito.

13. Plan de Capacitación y Transferencia de Conocimiento

El contratista deberá diseñar y ejecutar un plan de capacitación integral.

- **Capacitación Funcional**
  - **Objetivo:** Asegurar que los usuarios puedan operar la solución de forma autónoma.
  - **Contenido:**

- Uso de la Databricks App para solicitar los reportes, ver histórico de solicitudes finalizadas, navegar por su ayuda.
- Uso de archivos resultantes (html y csv),
- Uso de notebooks.

- **Entregables:**
  - Manual de Usuario.
  - Presentaciones y material de apoyo.
  - Sesiones de capacitación.
  - Videos o guías rápidas de referencia.
  - Archivos editables de los documentos generados.

- **Transferencia de Conocimiento Técnico (Equipo Técnico DIAN):**

  - **Objetivo:** Habilitar al equipo de la DIAN para mantener, operar y evolucionar la solución.
  - **Contenido:** Arquitectura detallada, explicación del código de los pipelines, modelo de datos de la capa Gold, proceso de despliegue (CI/CD).
  - **Entregables:**
    - Manuales Técnicos
    - Manuales de Operación.
    - Documentación de la arquitectura.
    - Archivos editables de diagramas y documentación generada.
    - Código fuente de cualquier artefacto generado.
    - Actas de reuniones.
    - Presentaciones y en general cualquier tipo de documento de apoyo para la transferencia de conocimiento.
    - Documento o página en herramienta de seguimiento dispuesta por la DIAN al proveedor para facilitar la navegación del paquete documental (como tabla de contenidos o mapa de navegación).

14. Entregables

No.	Fase	Descripción de Actividades Principales y detalle entregables
1	***Plan de trabajo Detallado***	Será responsabilidad del contratista proponer y posteriormente ejecutar, como parte de su oferta técnica, un plan de trabajo detallado. Dicho plan deberá incluir un cronograma de entregas que defina las fases, hitos y fechas para la implementación y presentación de cada uno de los entregables listados en esta sección y la siguiente. El cronograma deberá estar diseñado para permitir entregas incrementales de la solución, facilitando así la realización de pruebas y validaciones progresivas por parte de la DIAN. El seguimiento al cumplimiento de este cronograma se llevará a cabo en las mesas de trabajo periódicas (semanales) ya mencionadas, donde se formalizará la recepción de cada entregable conforme a la metodología y formatos establecidos en este Anexo. Para las mesas de trabajo o sesiones realizadas, Formatos de actas generadas.
2	***Especificaciones funcionales del PMV***	Registros en la herramienta de seguimiento azure DevOps completamente finalizados (work ítems tipo task, issue y epic en estado “Done” con sus respectivas evidencias vinculadas o adjuntas):

		Formato FT-IIT-2006 (Historias de usuario), Formato de entrega requerimientos FT-IIT-2008, Formatos de actas generadas
3	***Diseño técnico y maquetas***	Registros en la herramienta de seguimiento azure DevOps completamente finalizados (work ítems tipo task, issue y epic en estado “Done” con sus respectivas evidencias vinculadas o adjuntas): FT-IIT-2003 Modelo de diseño Elaboración de artefactos/diagramas de diseño requeridos por las HU de la(s) etapas en donde se aplique, entre otros: Diagramas de secuencia, diagramas de clases, diagramas de Entidad-Relación, diagrama de Despliegue. Formato de prototipado y testeo (FT-IIT-2712), Formatos de actas generadas El proveedor deberá entregar un dimensionamiento de la solución que entrega a la DIAN y las consideraciones que considere deban tenerse en cuenta para el crecimiento futuro por la utilización de la plataforma, debe incluir calculadora o estimación para dimensionar los costos de infraestructura, se deberá realizar para aprobación del diseño por parte de DIAN y se generará nueva versión para aprobación como requisito para solicitar cualquier promoción (pull request).
5	***Entrega de PMV App de Databricks con funcionalidad de solicitud y visualización de estado gestión solicitud tras aprobación de pruebas y validación de cumplimiento requisitos.***	App Databricks para registrar solicitudes de reportes tipificados funcionando Según los numerales 9. Reportes, 10. Requerimiento funcionales y no funcionales y 11. Reglas de negocio, Registros en la herramienta de seguimiento azure DevOps completamente finalizados (work ítems tipo task, issue y epic en estado “Done” con sus respectivas evidencias vinculadas o adjuntas): FT-IIT-2725 Plan de pruebas funcionales Evidencias de pruebas unitarias mediante reporte de cobertura de código u otras herramientas que el contratista disponga. Para cualquier conjunto de Pruebas se requiere como mínimo: Código fuente, Documentación sobre cómo se definieron los datos de prueba cuando se hayan usado conjuntos con tales características y evidencias de ejecuciones satisfactorias El proveedor debe entregar a DIAN todos los artefactos que le permitan a la DIAN repetir de manera autónoma la automatización de pruebas en ambiente que disponga la DIAN, las pruebas funcionales se deben automatizar. Para aquellos flujos y funcionalidades que no puedan ser automatizados se realizaran pruebas manuales. Estas pruebas incluirán la validación de la interfaz de usuario, la usabilidad, y la verificación de escenarios específicos de negocio que requieran intervención

		humana para su validación, las cuales deben estar documentadas y aprobadas en los formatos establecidos en el sistema de gestión de la Entidad (uso de DevOps de la DIAN). Siempre que se posible se deberán automatizar pruebas de regresión *Para la validación de requisitos no funcionales puede tomarse como inspiración: Manual de Fundamentos de ingeniería. Captura de requisitos no funcionales https://microsoft.github.io/code-with-engineering-playbook/design/design-patterns/non-functional-requirements-capture-guide/ Formato FT-IIT-2708 (acta para entrega de solución tecnológica) enfocada en monitoreo general que tendrá que realizarse una vez el producto inicie su fase de laboratorios de uso real en el periodo de la garantía.
6	***Entrega de ETLs mediante notebooks***	Workflows en el espacio de trabajo databricks dispuesto con sus notebooks y ejecuciones de prueba realizadas. Registros en la herramienta de seguimiento azure DevOps completamente finalizados (work ítems tipo task, issue y epic en estado “Done” con sus respectivas evidencias vinculadas o adjuntas): Documentación de las ETL. Documentación de las pruebas técnicas. Documentación de las pruebas funcionales. Aceptación de áreas de negocio. Según los numerales 9. Reportes, 10. Requerimiento funcionales y no funcionales y 11. Reglas de negocio Documentación sobre cómo se definieron los datos de prueba y evidencias de ejecuciones satisfactorias: El proveedor debe entregar a DIAN todos los artefactos que le permitan a la DIAN repetir de manera autónoma la automatización de pruebas en ambiente que disponga la DIAN. ***Para las ETL NO se requieren pruebas funcionales pero si se requiere validar calidad de datos.*** Para aquellos flujos y funcionalidades que no puedan ser automatizados se realizaran pruebas manuales., las cuales deben estar documentadas y aprobadas en los formatos establecidos en el sistema de gestión de la Entidad (uso de DevOps de la DIAN). Siempre que se posible se deberán automatizar pruebas de regresión *Para la validación de requisitos no funcionales puede tomarse como inspiración: Manual de Fundamentos de ingeniería. Captura de requisitos no funcionales

		https://microsoft.github.io/code-with-engineering-playbook/design/design-patterns/non-functional-requirements-capture-guide/ El proveedor una vez superado el desarrollo, deberá entregar el código fuente en el repositorio dispuesto por la DIAN y las configuraciones necesarias para realizar el despliegue continuo y el aprovisionamiento de los recursos, de acuerdo con los lineamientos o procedimientos establecidos por el equipo técnico. Finalmente, se registará en formato FT-IIT-2708 (acta para entrega de solución tecnológica) la administración o monitoreo general que tendrá que realizarse una vez el producto inicie su fase de laboratorios de uso real en el periodo de la garantía.
7	***Entrega de plantillas html para visualización de resultados que permitan a las ETL generar los reportes html reales.***	Archivos html originales de las plantillas base para la generación dinámica de reportes en las ETL cumpliendo lo indicado en los numerales 9. Reportes, 10. Requerimiento funcionales y no funcionales y 11. Reglas de negocio. Con los html se tendrá que validar con usuarios de negocio de las áreas que se considere podrán ser los principales usuarios finales, que representan reportes valiosos y de fácil lectura, la interacción javascript deberá ser validada en distintos navegadores y de requerirse ajuste en el diseño de su interfaz o en general se deba modificar la plantilla por no cumplir con las expectativas o metas definidas con el o los expertos UX que aporte el contratista, documentar tales labores en actas. Registros en la herramienta de seguimiento azure DevOps completamente finalizados (work ítems tipo task, issue y epic en estado “Done” con sus respectivas evidencias vinculadas o adjuntas): Formato FT-IIT-2006 (Historias de usuario), Formato FT-IIT-2428 Especificación Reporte
8	***Informe de Resultado de Pruebas***	Es un documento formal que consolida y resume los resultados de las pruebas unitarias y de integración, unitarias y de integración, del sistema end -to- end, de carga/stress y de aceptación de usuario (UAT) realizadas, este informe certifica que la solución tecnológica cumple con el numeral 12. Estrategias de calidad, pruebas y criterios de aceptación Debe contemplar lo documentado en el plan de pruebas y en general cualquier labor realizada según el tipo de prueba implementada: • De integración • De rendimiento/carga • De seguridad • Funcionales (Formato FT-IIT-2725) • Técnicas no funcionales (para validar requerimientos no funcionales)

		Para obtener mayor detalle sobre los criterios de aceptación mínimos en cualquier prueba, consultar los numerales correspondientes a tal temática dentro de los lineamientos de desarrollo de software DIAN v1.3: https://www.dian.gov.co/dian/Documents/Lineamientos-Desarrollo-SW-DIAN-V-1-3.pdf#search=FT%2DIIT%2D2708
9	***Certificados o actas de capacitación y cualquier documento generado en las transferencias de conocimiento.***	Son documentos que validan la formación impartida a los equipos funcionales y técnicos en un proyecto de desarrollo tecnológico. Sirven como prueba formal de que los usuarios y los equipos de soporte han recibido el conocimiento necesario para operar y mantener la nueva solución. Se debe aplicar una metodología participativa en las capacitaciones y sesiones de transferencia de conocimiento garantizando que: • Se iniciarán las sesiones con activadores como preguntas interactivas, dinámicas de reflexión o pequeños retos para enganchar a los participantes desde el inicio. • Las sesiones incluirán dinámicas interactivas, como grupos de trabajo, encuestas en vivo y sesiones de retroalimentación, para mantener el interés y fomentar la participación. • Para poder verificar el uso de la metodología, al finalizar cada sesión se debe medir la satisfacción y percepción de la audiencia mediante encuesta usando Microsoft forms desde una cuenta de funcionario DIAN, y los resultados serán parte de la documentación. Ver numeral 13.*Plan de Capacitación y Transferencia de Conocimiento*
10	***Paquete documental final entregado.***	Este paquete incluye: • Manuales Técnicos ▪ Manuales de Operación. ▪ Documentación de la arquitectura. ▪ Archivos editables de diagramas y documentación generada. ▪ Código fuente de cualquier artefacto generado. ▪ Actas de reuniones. ▪ Presentaciones y en general cualquier tipo de documento de apoyo para la transferencia de conocimiento. ▪ Documento o página en herramienta de seguimiento dispuesta por la DIAN al proveedor para facilitar la navegación del paquete documental (como tabla de contenidos o mapa de navegación).
11	***Acta de Cierre de Proyecto.***	Es un documento formal que marca la finalización oficial de un proyecto. Su propósito es certificar que se han completado todas las actividades y entregables planificados, y que el proyecto ha cumplido con sus objetivos iniciales, según lo establecido en el Acta de Constitución y el Plan de Proyecto. Este documento clave incluye varias secciones importantes: la Identificación del

		proyecto, que contiene los datos básicos; los Resultados y logros, que detallan los entregables y los resultados obtenidos; el Desempeño del proyecto, que resume el cumplimiento en términos de alcance, cronograma y presupuesto; las Lecciones aprendidas, donde se analizan los éxitos y desafíos; y finalmente la Aprobación formal, que se sella con las firmas del gerente del proyecto y otras partes interesadas, validando la finalización y aceptación del proyecto.

15. Factores de Ponderación de Calidad y Valor Agregado

Para la evaluación de las propuestas, además del cumplimiento de los requisitos obligatorios, se valorarán positivamente los siguientes elementos de calidad y valor agregado (es decir elementos que superan lo solicitado como producto base) que el contratista pueda ofrecer

15.1. Diseño e Implementación de dashboards de monitoreo técnico para observar la salud, rendimiento y costos de los pipelines. 16.5 Puntos

Se busca tener la visibilidad de la operación y facilitar la identificación de comportamientos no esperados de los pipelines de datos, permitiendo un monitoreo proactivo, para lo cual el proveedor debe realizar el diseño, desarrollo e implementación de un conjunto de tableros de monitoreo integrales que cubren exhaustivamente la salud (ej., errores, fallos), el rendimiento (ej., latencia, volumen procesado, tiempo de ejecución) y los costos (ej., consumo de recursos, optimización de gastos en la nube) de los pipelines implementados en la ejecución del presente proceso.

Las fuentes de datos para el monitoreo serán los propios pipelines y la infraestructura que los soporta, las herramientas a utilizar deben asegurar que no se requiera licenciamiento adicional para la DIAN, pueden o no ser open source (ej., Power BI, Grafana, Kibana, Azure Monitor), las fuentes de datos para el monitoreo y cómo se visualizarán las métricas clave para facilitar la toma de decisiones proactiva.

Debe Incluir ejemplos o mockups de los diseños propuestos e instrucciones de instalación, configuración o puesta a punto de las herramientas necesarias.

Es crucial destacar que la totalidad del desarrollo será sujeto a ***revisión y aprobación*** por parte de la supervisión del contrato de la DIAN. *En NINGÚN CASO se dará por entendido que un entregable es aprobado automáticamente sin pasar por las actividades que aseguren su calidad y entrega a satisfacción*.

15.2. Documento metodología de pruebas automatizadas (unitarias y de integración) para los pipelines de datos 14 Puntos.

Entrega de definición de un documento que detalle el proceso para realizar las pruebas automatizadas para cada etapa del pipeline (ingesta, transformación, carga), Documentación que indique el enfoque, las herramientas específicas para automatización con herramientas que no requieran licenciamiento o adquisición adicional (ej. PyTest, Great Expectations, DBT testing).

En el documento se incluye como mínimo: estrategia de cobertura de pruebas, Un plan claro para la ejecución y reporte de resultados enfocadas en asegurar la fiabilidad de los flujos de datos.

La propuesta debe demostrar un conocimiento profundo de automatización de las pruebas en procesamiento de datos, el cual será sujeto a ***revisión y aprobación*** por parte de la supervisión del contrato. *En NINGÚN CASO se dará por entendido que un entregable es aprobado automáticamente sin pasar por las actividades que aseguren su calidad y entrega a satisfacción*.

15.3. Implementación de un módulo en la app databricks para permitir al usuario descarga directa de sus archivos de reportes una vez fueron gestionados (funcionalidad que permitiría complementar las descargas desde unity catalog). 14 Puntos

Proveer un módulo en la aplicación (databricks app) que permita al usuario final la descarga directa de sus archivos de reportes generados a solicitudes registradas como otra funcionalidad del producto, de forma que todo el ciclo (desde la solicitud hasta la obtención del reporte final), se maneje dentro del aplicativo, lo que aumenta la coherencia y la facilidad de uso.

El usuario podrá visualizar el estado de sus solicitudes de reporte, desde "En proceso" hasta "Completado". Una vez que un reporte esté listo, además del estado completado se presentará un botón de descarga claro y visible debe habilitarse para que el usuario pueda obtener los archivos directamente (tanto el csv como el html).

Se debe contemplar un Diseño de Interfaz de Usuario (UI): que prioriza una experiencia intuitiva al ofrecer un panel de control personalizado donde los usuarios pueden monitorear el estado de sus solicitudes y descargar sus archivos directamente sin necesidad de acceder al volumen de unity catalog.

Lógica de la Aplicación (Backend): que garantiza la interacción con los volúmenes de Unity Catalog, la gestión segura de las solicitudes y la conexión con los procesos ETL que generan los reportes.
Aseguran la fiabilidad y el rendimiento del sistema, incluyendo una gestión robusta de errores y la optimización para descargas eficientes, incluso de archivos grandes.

Es crucial destacar que la totalidad del desarrollo será sujeto a ***revisión y aprobación*** por parte de la supervisión del contrato de la DIAN (*en NINGÚN CASO se dará por entendido que un entregable es aprobado automáticamente sin pasar por las actividades que aseguren su calidad y entrega a satisfacción*).

16. Metodología y documentación

Se registrará el proceso en azure DevOps, y el acceso es gestionado por la DIAN para asegurar que el proveedor pueda digitar su avance en las labores que van comprometiéndose en cada iteración.

Para cada reporte que se defina como posible elemento a solicitar mediante la solución que servirá como interfaz para tal fin, el proveedor o contratista deberá registrar su definición, características, validaciones mediante diversos formatos como los siguientes:

FT-IIT-2428 (Formato especificación reporte)
(https://diancolombia.sharepoint.com/:x:/s/diannetpruebas/Areas/EQcdQUhDFgVEtcxDic9kBSMBw8pgWDjRpgfDSznQUR6Iqw)

FT-IIT-2006 (Historias de usuario)

(https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/Ef4Ur3QHmnRFlbdixlni858BZbvFBABDJfPDG5glX6AVrg?e=ae8EPk)

*Una vez se hayan finalizado las especificaciones que se requieran, se recomienda diligenciar el formato de entrega requerimientos FT-IIT-2008 (*https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/EQVx7PinnwNItrztp_8t_t4BwRG79L4DLbnA9VmuFFDw_Q*)*

A lo largo del proceso, a medida que se generen prototipos y se validen módulos o entregables, el proveedor diligenciará el formato FT-IIT-2712 (prototipado) para registrar que se ha probado o validado satisfactoriamente: (https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/EZjDtR2ICF5Fgs5yfWlXNqABLsQAPHvythL6pfaPhzhv6w?e=FdpbyD)

Cuando se realicen reuniones o mesas de trabajo y sea requerido generar una ayuda de memoria o documentar la sesión, emplear el formato general de acta: (https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/ESozUDoaVINLiQBDGV_8F68BdHFgScuwiCDZv6AEtIIAsQ?e=q3VZn6)

Una vez se haya implementado la solución y se considere que el producto está finalizado, se tendrán que validar los resultados y registrar la satisfacción de la totalidad de entregables mediante el formato acta para entrega de solución tecnológica FT-IIT-2708: (https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/EVSuZRn0Uu9Os-_QAFHi2VQBV-0m-MLaJNsZrgzAT_Hgsw?e=fSkIDS)

Si llegase a requerirse un acta de terminación de mutuo acuerdo para contrato, diligenciar el formato FT-ADF-2092 (https://diancolombia.sharepoint.com/:w:/s/diannetpruebas/Areas/Edakeu-AI6lGhYVhybf0J2MB3_fJInLZYaHXZ9KNXtxv1g?e=AgofB5)

Finalmente, el proveedor deberá entregar la documentación de manuales de uso, manuales técnicos y/o cualquier otro elemento documental generado.

En NINGÚN CASO se dará por entendido que un entregable es aprobado automáticamente sin pasar por las actividades que aseguren su calidad y entrega a satisfacción.

17. Cumplimiento Normativo:

La solución debe adherirse a todas las leyes, regulaciones y políticas internas de la DIAN relacionadas con la gestión, protección y privacidad de datos, especialmente en lo que respecta a la información sensible y confidencial. Se debe proporcionar evidencia del cumplimiento de las políticas de seguridad de la información y privacidad de datos de la DIAN, incluyendo la gestión de datos sensibles de acuerdo con la normatividad vigente.

- **Garantía de Estabilidad y Calidad del Producto**: El contratista está obligado a proporcionar una garantía integral de un (1) año sobre la calidad, estabilidad y correcto funcionamiento de la solución implementada. Esta garantía se activará una vez finalizado y aceptado el contrato y estará amparada por la póliza de cumplimiento.

    - **Alcance de la Cobertura**: La garantía cubre la corrección sin costo adicional de cualquier defecto de implementación, error de software (bug), o vicio oculto atribuible al desarrollo realizado por el contratista. Esto incluye, pero no se limita a, fallos en la lógica de los pipelines, inconsistencias en los datos generados por errores de codificación, problemas

de rendimiento directamente relacionados con un diseño deficiente, o cualquier otra desviación funcional respecto a lo especificado en este anexo técnico.

- **Exclusiones**: La garantía no cubrirá incidencias derivadas de cambios en la infraestructura base de la DIAN no notificados, modificaciones en los datos fuente no contempladas, o alteraciones del código realizadas por personal ajeno al contratista después de la entrega.

- **Proceso de Notificación y Soporte**: La DIAN notificará formalmente cualquier incidencia cubierta por la garantía, y el contratista deberá iniciar el diagnóstico y proponer un plan de acción dentro de los plazos de respuesta que se acuerden en el marco del contrato.

- **Soporte técnico**. Mediante el cual se brinda apoyo o asistencia técnica a los especialistas de la Entidad ante incidentes técnicos que afecten la operación, replicación de datos y *Fail Back*, con el fin de detectar la causa raíz de las problemáticas detectados resolver *técnicamente el incidente. Este soporte se debe brindar a demanda como parte de la garantía del producto. El proveedor deberá alinearse con el modelo de soporte de la DIAN (modelo operativo y herramientas), es decir, su soporte deberá organizarse para que procedimentalmente opere con el funcionamiento de los equipos técnicos de la DIAN. El mantenimiento y soporte será para el producto desarrollado, sin embargo, si el desarrollo realizado afecta o modifica parte de componentes existentes en funcionamiento, también deberá darse soporte a tales componentes.*

Nota aclaratoria:

La DIAN podrá realizar un fork del producto (crear otra versión de la solución) mientras la garantía está vigente si requiere realizar ajustes evolutivos al producto y probarlos de manera independiente en un ambiente sandbox o de desarrollo, bajo análisis y aprobación de todas las áreas involucradas quienes asumirían la responsabilidad de tener en simultáneo dos versiones de la aplicación, que se tratarían como “Versión productiva y versión alpha”.

Si se realizaran tales labores en el término de la garantía, y se ha alcanzado madurez en una nueva característica (por ejemplo un nuevo reporte), podrá informarse al equipo contratista para que analice el aumento evolutivo, y determinar si considera viable su inclusión (merge) en el producto vigente en producción y podrá también indicar si genera alguna exclusión específica de la garantía vigente. Tras la determinación del impacto las partes podrán determinar si se procede a la fusión en la rama de código fuente productiva (master).

18. Otros aspectos a considerar

- Cláusula de Cesión de Derechos Patrimoniales de Propiedad Intelectual sobre el Software Desarrollado. Todos los desarrollos realizados en ejecución del contrato, incluyendo el código fuente, código objeto, scripts, algoritmos, diseños, configuraciones, manuales y demás elementos que resulten del presente contrato, incluidos los que se generen durante el análisis, diseño, desarrollo, implementación, prueba y puesta en funcionamiento de la solución tecnológica contratada serán de propiedad exclusiva de la DIAN.
  El proveedor cede a título gratuito y definitivo los derechos patrimoniales de autor, esta cesión incluye, pero no se limita a, los derechos de reproducción, transformación, distribución, comunicación pública y cualquier otra modalidad de explotación, conforme a lo dispuesto en la Ley 23 de 1982, la Decisión Andina 351 de 1993 y demás normas que las modifiquen o sustituyan, sin limitación territorial ni temporal.

- Cualquier componente de terceros, deberá contar con licencia para uso a perpetuidad y aprobación previa de la DIAN. El proveedor deberá entregar el código fuente y suscribir el acta de cesión antes de finalizar el contrato.
- El proveedor deberá hacer entrega formal al equipo de tecnología de la DIAN, explicando todo el desarrollo y dejando por escrito lo necesario para adoptar su evolución y mantenimiento posterior.
- SPD representará el apoyo técnico y en ningún caso realiza labores relacionadas a la supervisión del proceso.
- La DIAN deberá revisar y aprobar todo desarrollo y documentación/entregables.
- El proveedor deberá entregar todos los manuales de desarrollo y de uso para el desarrollo completo.
- El proveedor deberá actualizar los manuales de uso vigentes en la entidad para todo lo ajustado por él.
- Si la DIAN no cuenta con documentación técnica y el proveedor deberá revisar el código y realizar la ingeniería inversa con el fin de adquirir el conocimiento necesario para adelantar la implementación del proyecto.
- Si el proveedor define algún cambio de tecnología de datos o su diseño necesita algún tratamiento de datos, deberá asumirlo como parte del contrato.
- Cuando se trate de fábricas el costo de las suscripciones de las herramientas de desarrollo deben ser asumidos por ellos durante el proyecto y la garantía. Las herramientas de desarrollo deben ser aprobadas por la DIAN.

Firmas:

VALERO VASQUEZ LUIS HERNANDO Firmado digitalmente por VALERO VASQUEZ LUIS HERNANDO Fecha: 2025.10.24 14:21:15 -05'00' LUIS HERNANDO VALERO VASQUEZ *Subdirector de Factura Electrónica y Soluciones Operativas*	OMAR ROJAS GARCIA *Subdirector de Procesamiento de Datos (A)*

Elaboró	F. Gustavo Sierra Hernández	Gestor II	Subdirección Procesamiento de Datos.
Revisó	Carlos Eduardo Vargas Gallo	Inspector III	Subdirección Factura Electrónica y Soluciones Operativas
Aprobó	Omar Rojas Garcia	Subdirector (A)	Subdirección Procesamiento de Datos
	Luis Hernando Valero Vásquez	Subdirector	Subdirección Factura Electrónica y Soluciones Operativas