TABLA DE CONTENIDO

1. INTRODUCCIÓN 2

2. OBJETIVO 2

3. ALCANCE 2

4. DESARROLLO 2

4.1. REQUERIMIENTOS FUNCIONALES 2

4.2. REQUERIMIENTOS NO FUNCIONALES 8

4.3. PRIORIZACION DE ATRIBUTOS DE CALIDAD 12

4.4. TÁCTICAS DE ARQUITECTURA 13

4.5. VERIFICACIÓN DE TÁCTICAS DE ARQUITECTURA 14

4.6. ARQUITECTURA ACTUAL Y DE REFERENCIA 15

5. CONTROL DE CAMBIOS 17

CLASIFICADA

1. INTRODUCCIÓN

Para atender la necesidad del ICBF de contar con información unificada y en tiempo real para la toma de decisiones, el seguimiento y mejora en la prestación de sus servicios, se plantea como solución la implementación de proyectos de analítica avanzada y la generación de indicadores. Este documento referencia la arquitectura que deberá aplicarse en adelante para los proyectos de analítica que se lleven a cabo en la entidad.

La arquitectura de datos agrupa las especificaciones de requisitos de datos, guía la integración y el control de los activos de datos, apoya las necesidades de información y permite la estandarización e integración de datos.

Se presentan los requerimientos funcionales, no funcionales, recomendaciones y mejores prácticas que deben tenerse en cuenta, priorización de los requerimientos no funcionales, estableciendo las principales preocupaciones en desempeño, disponibilidad, interoperabilidad y seguridad. Además de la validación de dichos lineamientos para cualquier solución que involucre analítica avanzada. Además, se involucra el estado actual de la arquitectura del ICBF y finalmente la arquitectura propuesta a ser implementada en los proyectos concernientes.

2. OBJETIVO

Adoptar los estándares de arquitectura de referencia para todos los proyectos concernientes a la analítica avanzada, proporcionando un vocabulario común de estándares de negocio y requisitos de datos estratégicos, alineando los proyectos analíticos con los diseños de integración, estrategia empresarial y arquitectura del negocio.

3. ALCANCE

Aplica en el nivel de la Sede Nacional.

4. DESARROLLO

4.1. REQUERIMIENTOS FUNCIONALES

Los siguientes requerimientos funcionales aplicables a todos los proyectos de analítica avanzada, establecen los pilares deseables para la construcción del proyecto, y apoyan en la generación de requerimientos, llevándolos finalmente a las salidas esperadas para estos.

- Estudio y comprensión de datos.

CLASIFICADA

Documentar, a partir de las fuentes de información analizadas, las metodologías que apliquen para la implementación de los diferentes modelos de datos, en los marcos de Analítica de datos de la entidad identificando los siguientes factores.

4.1.1. Identificación fuentes de información

Se debe determinar el acceso a la información y su ubicación. Estas deben contar con protocolos de seguridad y latencia determinados en los Requerimientos No Funcionales de seguridad.

4.1.2 Identificación de datos estructurados, semiestructurados, y no estructurados.

Determinar de las fuentes de datos cuáles están categorizadas en los tipos estructurados, semiestructurados y no estructurados.

4.1.3. Definición de metodología de consolidación y procesamiento.

Establecer la metodología de consolidación y procesamiento de la información que sirva como insumo para los cálculos de los datos e indicadores definidos.

Para proyectos de Machine Learning e Inteligencia Artificial se debe tener en cuenta lo siguiente.

Proceso	Descripción
Ingesta de datos	Debe permitir gestionar fuentes de datos y realizar cargas desde ambientes de Big Data tipo Data Lake. Debe permitir el perfilamiento de la calidad de datos sin salir del ambiente Big Data. Durante el proceso de integración se debe permitir desarrollar tareas orientadas a mejorar y mantener la calidad de datos como filtros, búsquedas, cálculos de estadísticas, entre otros. Deberá permitir un análisis de impacto en las tareas de integración de datos. Debe permitir ejecutar los procesos a través de una interfaz gráfica, sin necesidad de manipulación, compilación o generación de código por parte del usuario final. Debe permitir la generación de llaves subrogadas retenidas y naturales. Debe ofrecer funciones como filtros, búsqueda de tablas o archivos, agregación, cálculos matemáticos y estadísticos. Debe permitir mantenimiento (corrección, inclusión, eliminación) de registros utilizando el flujo de aprobaciones. Debe permitir la interrupción de un proceso durante su ejecución y garantizar la integridad de los datos hasta el último punto de control. Debe permitir el reinicio del proceso desde el punto de interrupción. Debe facilitar la entrada rápida en producción a través de una interfaz "*drag and drop*" y a través de una interfaz de código.

CLASIFICADA

Proceso	Descripción
Analítica Avanzada	Debe permitir construir modelos estadísticos lineales (Regresión Lineal, Regresión logística, modelos lineales generalizados, Regresión Logit, Regresión Probit, etc.) Debe permitir construir modelos de árboles (Árbol de Decisión, Bosque Aleatorio, Gradient Boosting, etc.) Debe permitir construir Modelos de Redes neuronales Artificiales. Debe permitir construir Modelos de Máquinas de soporte vectorial. Debe permitir construir Modelos para detección de anomalías. Debe permitir la construcción de modelos analíticos de aprendizaje no supervisado. Debe permitir el ajuste óptimo de híper-parámetros de modelos con técnicas tradicionales y avanzadas de forma automatizada (Grid Search, Algoritmos genéticos, métodos de inteligencia artificial). Debe estar en capacidad de construir modelos analíticos desde una interfaz “click-and-point” así como desde una interfaz de programación con lenguajes propietarios y open source. Debe permitir la generación automática de resúmenes de comparación de modelos a través de gráficos de umbrales, curvas ROC, estadísticas de concordancia y tablas de clasificación errónea para uno o más modelos. Debe permitir evaluar interactivamente los umbrales de corte para diferentes percentiles. Debe permitir exportar modelos como códigos, permitiendo que los modelos se integren con otras aplicaciones. Debe incluir análisis de texto, incluyendo nube de texto con análisis de "peso", así como análisis de sentimientos en documentos o en una colección de documentos. Debe conectarse con las redes sociales para explorar temas y comprender el sentimiento de los comentarios en Facebook, Twitter, Google Analytics y Youtube. Debe permitir análisis contextual: descubrimiento, categorización y sentimientos utilizando Conversión de documentos. Debe incluir análisis, comprensión previa del texto, identificación automática de la entidad. Debe permitir análisis de términos - Validación de ortografía, búsqueda avanzada de texto y filtros. Debe permitir crear nube de palabras con información no estructurada. Debe permitir la identificación automática de temas de documentos y análisis de categoría. Debe estar habilitado para el análisis de sentimiento de documentos. Debe permitir el procesamiento del lenguaje natural (PLN) y extracción contextual. Debe analizar y convertir texto en representaciones formales para el procesamiento y comprensión de lectura. Debe permitir el análisis de textos y contextos con técnicas de: tokenización, segmentación de palabras y frases, *stemming*, descomposición en bolsas de términos, y reconocimiento de entidades nombradas. Debe poseer las funcionalidades arriba en su mayor parte por una interfaz *“drag and drop”* para que mismos usuarios que no saben programar puedan utilizar la misma.

CLASIFICADA

Proceso	Descripción
	Debe estar en capacidad de documentar los parámetros, variables, características, y conjuntos de datos usados para construir un modelo, en un repositorio único de registro a través de proyectos. Debe tener la capacidad para hacer seguimiento y versionar modelos. Debe contar con un repositorio centralizado donde se puedan almacenar y compartir modelos analíticos, incluyendo sus códigos fuente, y códigos para reentrenamiento y puesta en producción. Debe tener la capacidad de monitorear el desempeño de modelos analíticos en el tiempo de forma automática. Debe permitir la validación de los modelos antes de exportarlos a producción. Debe contar con la capacidad de automatizar el reentrenamiento de los modelos analíticos. Debe tener la capacidad para construir flujos de trabajo que permitan estructurar procesos para monitorear el desempeño de los modelos analíticos, disparando alertas si un modelo en producción se descalibra y tomando decisiones al respecto. Debe contar con la capacidad para desplegar modelos analíticos en tiempo real, a través de APIs, en Batch y en Bases de Datos. Debe permitir la elaboración de informes sobre el rendimiento de los modelos durante el ciclo de vida. Debe contar con la capacidad de incluir en el ciclo de gestión de modelos, modelos de múltiples plataformas y modelos construidos en lenguajes open-source. Debe estar habilitado para la programación de tareas de scoring para que se ejecuten en horas y fechas específicas. Debe permitir importar y exportar códigos de modelo en PMML.
Reglas de Negocio	Debe tener una interfaz gráfica de usuario tipo drag and drop que permita construir gráficamente las reglas de negocio. Debe permitir la integración de modelos de inteligencia artificial dentro del flujo de decisión. Debe permitir la integración con repositorios de datos. Debe permitir el uso de reglas SI-ENTONCES-O para la ejecución automática de decisiones y reglas. Debe permitir parametrizar las reglas, decisiones y modelos dentro de la misma interfaz. Debe contar con control de versiones para hacer seguimiento y mantener el gobierno durante la implementación de las reglas. Debe integrarse con tablas de búsqueda para ejecutar búsquedas y verificaciones dentro de la lógica de implementación de las reglas.

4.1.4 Definición de modelo de calidad de los datos y depuración de datos.

Debe establecerse la metodología para la aplicación de modelos de calidad de datos y procesos de depuración de datos, estos deben garantizar unicidad, confiabilidad, escalabilidad, seguridad.

CLASIFICADA

4.1.5. Lineamientos de gestión de datos.

Debe existir una estructuración de los modelos de gestión jerárquica para las métricas de indicadores a nivel operativo, táctico y estratégico.

4.1.6. Estrategia de implementación.

Este proyecto se enmarca en las buenas prácticas del PMI para lo cual el contratista, deberá desarrollar los siguientes grupos de procesos:

- Planeación
    - Reunión de inicio (apertura del proyecto).
    - Entendimiento del problema y datos actuales.
    - Entregables de planeación
- Ejecución
- Seguimiento y Control
- Cierre y finalización del proyecto

4.1.7. Entregables técnicos

Los entregables técnicos deben contemplar lo siguiente.

- Análisis de datos y caracterización: El cual debe contemplar el modelo de entrada de datos de fuentes internas y/o externas, el modelo de almacenamiento de datos de fuentes internas y/o externas y el modelo de integración.
- Modelado de información: Debe contener modelo de vista funcional (presentación de los datos), vista estructural (componentes a alto nivel), vista de implementación (flujo de información desde la fuente a la presentación y modelo de calidad de datos).
- Documento de diseño del modelo y los componentes de datos.
- Descripción de los componentes o artefactos de extracción, consolidación, procesamiento, calculo, gestión histórica y proyección de la información. procedimientos almacenados a ser modificados.
- Documento de diseño de interfaz de los reportes y tableros de control a presentar por producto.
- Documento de administración de componentes del modelo de datos (como mínimo procedimiento de seguridad y procedimiento de respaldo).
- Modelo de vista de despliegue (como mínimo diagrama de despliegue).
- Diccionario de datos.

Nota: Los entregables que apliquen al proyecto, se definen en la reunión de inicio (apertura del proyecto).

CLASIFICADA

4.1.8. Modelo de seguridad y segmentación de datos.

Debe contener el Modelo para el acceso a fuentes de información. Modelo jerárquico y segmentado para presentación de información (Roles por visualización y limitantes).

4.1.9. Obtención y presentación de datos.

Debe realizarse la materialización en Scripts, componentes y/o artefactos para la gestión de la información a través de modelos de Analítica. El código implementado y/o entregado debe estar bajo supervisión del ICBF en la herramienta establecida para tal fin. Además de ello, la presentación de los datos debe estar acorde a los colores corporativos de la entidad.

4.1.10. Modelos de búsqueda, extracción, consolidación y procesamiento.

Deben generarse Scripts, componentes, y/o artefactos para la gestión de búsqueda y extracción de información, consolidación de información y procesamiento de las diferentes líneas del modelo.

4.1.11. Despliegue de soluciones de Analítica avanzada.

Debe generarse scripts de código fuente documentado, control de versiones de la solución, componentes de instalación, scripts y ejecución de scripts de pruebas unitarias, manuales de instalación, operación y configuración. Instalación en ambientes de prueba, casos de prueba funcionales y no funcionales. Informe de pruebas que involucre las siguientes pruebas en los casos que aplique.

- Pruebas de funcionalidad
- Pruebas de configuración
- Pruebas de seguridad y de control de acceso.
- Pruebas de interfaz de usuario.
- Registro de incidentes o defectos, su tratamiento y tiempo de solución.
- Pruebas de base de datos e integridad de datos.
- Pruebas de concurrencia.
- Pruebas de desempeño.
- Pruebas de carga
- Pruebas de recuperación de fallas.
- Pruebas de servicios de interoperabilidad.
- Pruebas de aceptación

Estas pruebas deben cumplir con las especificaciones técnicas y funcionales definidas por parte del ICBF.

CLASIFICADA

4.1.12. Aprovisionamiento de infraestructura.

Debe contemplarse la entrega de Documento de arquitectura de despliegue con definiciones técnicas para la implementación de la solución.

En el despliegue de la funcionalidad también debe tenerse en cuenta.

- Formato de requerimiento de cambios (RFC).
- Documento protocolo de soporte.
- Formatos con asistencia a capacitación por parte de los usuarios funcionales y transferencia de conocimiento a enlaces técnicos de los indicadores implementados en la etapa de operación.
- Actas y lista de asistencia de reuniones realizadas.

4.2. REQUERIMIENTOS NO FUNCIONALES

Los siguientes requerimientos no funcionales constituyen los lineamientos básicos que debe tener una solución de analítica, los lineamentos básicos hacen énfasis en la parte técnica que se debe tener en cuenta para llevar a cabo una salida exitosa de un producto.

- **Desempeño**

ID	RNF-D1
Nombre del Requerimiento	Procesamiento paralelo en memoria distribuido
Descripción del Requerimiento	El sistema debe permitir, el procesamiento paralelo en memoria con el objetivo de obtener escalabilidad horizontal y vertical.

ID	RNF-D2
Nombre del Requerimiento	Tolerancia a fallos
Descripción del Requerimiento	El sistema debe ser capaz de responder a escenarios en los que se presenten fallos, por lo cual, se sugiere (según sea el caso) implementar balanceadores de carga dando completa disponibilidad de los servicios que presenten fallas. También se sugiere la implementación de servicios en estado Activo-Pasivo según sea el caso.

ID	RNF-D3
Nombre del Requerimiento	Tiempos de respuesta óptimos
Descripción del Requerimiento	Los tiempos de respuesta en los diferentes escenarios no debe superar los 3 segundos, ya que el usuario reaccionaría de manera negativa.

CLASIFICADA

- **Disponibilidad**

ID	RNF-DS1
Nombre del Requerimiento	Alta disponibilidad
Descripción del Requerimiento	El sistema debe contar con completa disponibilidad debido a su alto uso por parte de todas las aplicaciones, debe aplicar una disponibilidad del 99.5%, y de ser posible, las 44 horas de indisponibilidad al año puedan aplicarse fuera de horarios de alta demanda.

- **Auditabilidad**

ID	RNF-A1
Nombre del Requerimiento	Gobernabilidad del sistema
Descripción del Requerimiento	Para la entrega por parte del proveedor de analítica avanzada. La solución debe permitir la administración y orquestación desde una consola centralizada, tanto de los modelos como de los datos.

ID	RNF-A2
Nombre del Requerimiento	Control de acceso y auditoría
Descripción del Requerimiento	Todos los entregables de analítica deben llevar registro en el sistema de auditoría del producto.

- **Mantenibilidad**

ID	RNF – M1
Nombre del Requerimiento	Gestión del versionamiento
Descripción del Requerimiento	Uso de la herramienta Visual Studio Team Services (VSTS), para la implementación de la gestión del versionamiento.

ID	RNF – M2
Nombre del Requerimiento	Despliegue usando DevOps
Descripción del Requerimiento	El despliegue debe garantizarse a través de un pipeline estructurado usando esta metodología de implementación.

ID	RNF-M3
Nombre del Requerimiento	Parametrización del modelo analítico
Descripción del Requerimiento	El modelo analítico avanzado de inteligencia artificial debe estar basado en algoritmos de caja negra, como: Random Forest, Gradient Boosting, Redes Neuronales, entre otros.

CLASIFICADA

- **Usabilidad.**

ID	RNF-U1
Nombre del Requerimiento	Visualización de tableros de control
Descripción del Requerimiento	Se deberá construir información consolidada o desagregada de seguimiento y monitoreo. Este debe ser implementado en la herramienta PowerBI para su presentación.

ID	RNF—U2
Nombre del Requerimiento	Nombres en procedimientos almacenados, funciones, tablas, vistas, y demás artefactos de base de datos
Descripción del Requerimiento	Los nombres de procedimientos almacenados, funciones, tablas, vistas, y demás artefactos de base de datos, debe cumplir los estándares proporcionados por el ICBF dada en común acuerdo en la reunión de inicio del proyecto. También se pueden encontrar en la “G1.GTI Guía de Estándares de Desarrollo y Arquitectura de Sistemas de Información v9”

ID	RNF-8
Nombre del Requerimiento	Interfaz drag-drop
Descripción del Requerimiento	Se debe facilitar la entrada rápida en producción a través de una interfaz “*drag and drop*” y a través de una interfaz de código.

- **Interoperabilidad**

ID	RNF-I1
Nombre del Requerimiento	Interoperabilidad con otros sistemas
Descripción del Requerimiento	La solución debe permitir la interoperabilidad con el Lago de datos, y Bases de Datos del ICBF. Esta debe ser a través de servicios web o interfaces dispuestas para exposición de servicios. No está permitida la conexión a través de DB Links, Link Server, o alguna otra tecnología que realice conexiones directamente entre bases de datos.

ID	RNF-I2
Nombre del Requerimiento	Solución en Cloud
Descripción del Requerimiento	El sistema debe llevar su información a la plataforma actual de la entidad, la cual es Azure, en acompañamiento con el proveedor de servicios del Instituto Colombiano de Bienestar Familiar.

ID	RNF-I3
Nombre del Requerimiento	Lenguajes para análisis de datos
Descripción del Requerimiento	Modelamiento en lenguajes de análisis de datos usando el lenguaje Python, R, y SQL.

CLASIFICADA

- **Seguridad**

ID	RNF—S1
Nombre del Requerimiento	Autenticación y autorización
Descripción del Requerimiento	La solución debe estar en capacidad de integrarse con los sistemas de autenticación (Directorio Activo) y autorización del Instituto Colombiano de Bienestar Familiar. Adicionalmente, debe garantizar la seguridad en la transferencia de los datos *at-rest.*

ID	RNF—S2
Nombre del Requerimiento	Segmentación de visualización
Descripción del Requerimiento	La solución debe permitir la autorización o limitación de permisos dependiendo de los roles proporcionados por el Instituto Colombiano de Bienestar Familiar.

ID	RNF—S3
Nombre del Requerimiento	Seguridad de equipos y portátiles.
Descripción del Requerimiento	Los equipos de escritorio y portátiles que se utilicen en la ejecución del contrato deben contar con software licenciado y de seguridad. Estos equipos serán revisados por la Subdirección de Recursos Tecnológicos del ICBF quien dará el aval para ingreso y privilegios de acceso a la red de ICBF de manera controlada y segura.

ID	RNF—S4
Nombre del Requerimiento	Transferencia segura
Descripción del Requerimiento	El sistema debe transferir los datos de manera segura a través del protocolo de aplicación HTTPS.

ID	RNF—S5
Nombre del Requerimiento	Desarrollo de código seguro
Descripción del Requerimiento	Cumplimiento de calidad de código seguro alineado al TOP 10 de OWASP

ID	RNF—S6
Nombre del Requerimiento	Uso de encripción y desencripción de contraseñas
Descripción del Requerimiento	En el uso de componentes en Azure debe usarse Key Vault para almacenamiento de llaves teniendo en cuenta que estas deben ser encriptadas en el aplicativo. En el caso de componentes on-premise las cadenas de conexión en ambientes productivos deben encriptarse.

ID	RNF—S7
Nombre del Requerimiento	Firma de archivos

CLASIFICADA

Descripción del Requerimiento	Todos los archivos deben ser firmados para su entrega en el repositorio documental.

ID	RNF – S8
Nombre del Requerimiento	Cuentas de dominio para acceso a base de datos
Descripción del Requerimiento	Debe garantizarse el acceso a base de datos mediante cuentas de dominio y no por usuario y contraseña, estas cuentas deben ser proporcionadas en común acuerdo con el ICBF.

4.3. PRIORIZACION DE ATRIBUTOS DE CALIDAD

Las posibles calificaciones son: H: High; M: Medium; L: Low[1]. El resultado de esta priorización permite determinar los escenarios con prioridad alta, media y baja. Se hace énfasis en el desempeño, la disponibilidad, y la seguridad de los requerimientos no funcionales.

A continuación, se presenta el orden de priorización:

Atributo de Calidad	ID	REQUERIMIENTO NO FUNCIONAL	PRIORIDAD
Desempeño	RNF-D3	Tiempos de respuesta óptimos	H
Disponibilidad	RNF-DS1	Alta disponibilidad	H
Desempeño	RNF-D2	Tolerancia a fallos	H
Desempeño	RNF-D1	Procesamiento paralelo en memoria distribuido	H
Interoperabilidad	RNF-I1	Interoperabilidad con otros sistemas	H
Seguridad	RNF—S1	Autenticación y autorización	H
Seguridad	RNF—S2	Segmentación de visualización	H
Seguridad	RNF—S5	Desarrollo de código seguro	H
Interoperabilidad	RNF-I3	Lenguajes para análisis de datos	M
Mantenibilidad	RNF – M1	Gestión del versionamiento	M
Auditabilidad	RNF-A2	Control de acceso y auditoría	M
Interoperabilidad	RNF-I2	Solución en Cloud	M
Seguridad	RNF—S4	Transferencia segura	M
Seguridad	RNF—S6	Uso de encripción y desencripción de contraseñas	M
Seguridad	RNF – S8	Cuentas de dominio para acceso a base de datos	M
Seguridad	RNF—S7	Firma de archivos	L
Seguridad	RNF—S3	Seguridad de equipos y portátiles.	L
Auditabilidad	RNF-A2	Control de acceso y auditoría	L
Auditabilidad	RNF-A1	Gobernabilidad del sistema	L

[1] Utility Trees. Consultado el 07 de Marzo de 2014. Disponible en: http://www.sei.cmu.edu/reports/00tr004.pdf

CLASIFICADA

4.4. TÁCTICAS DE ARQUITECTURA

A continuación, se enumeran las tácticas de arquitectura que podrán ser seleccionadas para atender los atributos de calidad definidos.

Atributo de Calidad	Táctica de Arquitectura	Patrón de Arquitectura	Descripción	Tecnología de Referencia
Seguridad Requerimientos Atendidos Se cubre el 100% de los requerimientos no funcionales de Seguridad	Resistir Ataques	Autenticación de Usuarios Autorización de Servicios	Los usuarios deben autenticarse a través de un usuario y contraseña para poder hacer uso del sistema de información.	Integración con el componente de autorización del ICBF
	No repudio	Auditoria de eventos de aplicación y base de datos	Se registran todos los eventos del sistema	Integración con componentes de auditoria
	Resistir Ataques	Limitar la exposición	Utilización de herramientas Firewall para bloquear el acceso a los recursos del sistema. Los servidores de base de datos pueden estar en una zona militarizada, solo accesible desde la red interna	Top 10 OWASP.
Desempeño	Aumento de velocidad	Verificación de procedimientos almacenados	Debe realizarse verificación de procedimientos almacenados, donde las consultas eviten hacer uso de consultas en tablas completas y uso de cursores	Mejores prácticas en SQL
	Cargue de componentes	Verificación de páginas web y tableros de control	Debe garantizarse cargar únicamente los componentes necesarios con imágenes de buena resolución, es recomendable el tamaño de la página a cargar en promedio de 1,7Mb	Recomendaciones páginas web.
Disponibilidad	Alta disponibilidad de infraestructura	Redundancia de recursos en equipos alternos	Debe garantizarse la redundancia en equipos alternos (o regiones) dependiendo a lo que corresponda	Recomendaciones CISCO disponibilidad de sitios
	Alta disponibilidad de los datos	Activo/Pasivo Activo/Activo Mirroring	Debe garantizarse la alta disponibilidad de los datos ya sea duplicando su data a través de técnicas de respaldo.	Recomendaciones alta disponibilidad de Bases de datos

CLASIFICADA

Atributo de Calidad	Táctica de Arquitectura	Patrón de Arquitectura	Descripción	Tecnología de Referencia
Auditabilidad	Registros de auditoría en Cloud	Logging	Configuración de registros de auditoría	Recomendaciones de auditoría en Cloud.
		Privilegio mínimo	Aseguramiento de aplicación de controles adecuados (IAM)	
		Comprensión de registros	Opciones para ver los registros de auditoría y entendimiento de estos por administrador del sistema.	
	Registros de auditoría On-Premise	Seguridad en Logs	Establecer correcta infraestructura en manejo de logs.	Guía para la gestión de la seguridad de logs en sistemas de información dada en guía de mejores prácticas del MinTic
Usabilidad	Tableros de control y reportes	Intención de la información	Conocer propósito y audiencia de la información tienen una intención clara y a un público en específico	Mejores prácticas en usabilidad
		Ubicación	Aprovechar el lugar más visto (generalmente a la izquierda de la página / tablero o sección).	
		Ajuste de tamaño	Tableros deben poder ajustarse a los diferentes tamaños de los dispositivos.	
		Número de vistas	Evitar el uso de información concurrida o con bastantes datos	
		Filtrado	Uso de filtros para diferentes opciones de visualización	
Interoperabilidad	Entre sistemas de información	Intercambio simple	Solo lo necesario en la operación	Marco de interoperabilidad de gobierno de MinTic
		Eficiencia	Debe ser funcional para las partes involucradas	

4.5. VERIFICACIÓN DE TÁCTICAS DE ARQUITECTURA

Se ilustra una matriz que muestra la relación entre las tácticas de arquitectura seleccionadas y los respectivos escenarios y atributos de calidad. De esta forma se logra verificar que la arquitectura este completa y balanceada, verificando que no haya escenarios sin tácticas, que no haya tácticas repetidas o que estén en conflicto.

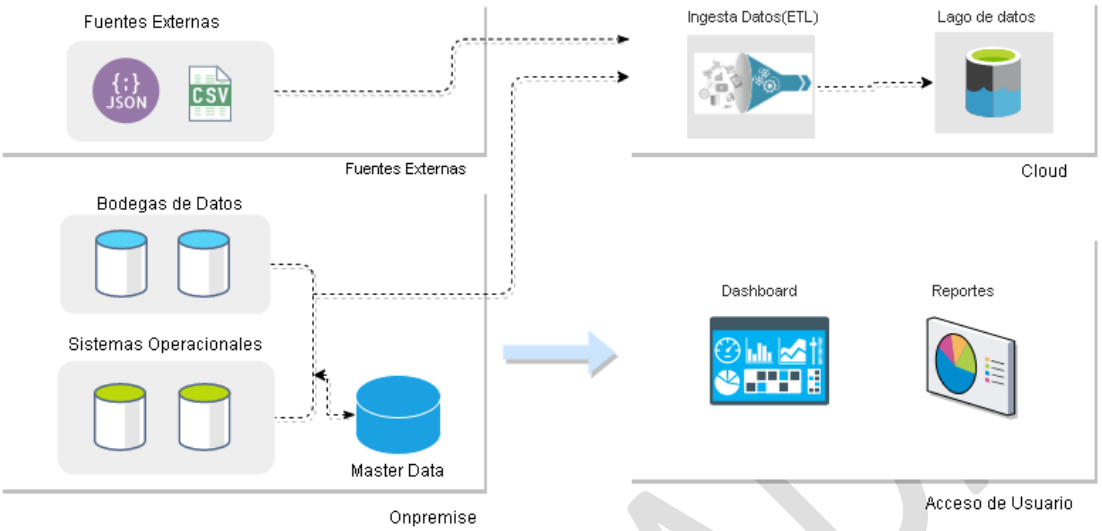
CLASIFICADA

ATRIBUTO DE CALIDAD	ESCENARIO DE CALIDAD	TÁCTICA DE ARQUITECTURA	CRITERIOS DE SELECCIÓN
Seguridad	Autenticación Autorización TOP 10 OWAS	Resistir ataques	La autenticación para usuarios externos e internos contra el módulo de autenticación que provee el ICBF que se incorporara en el sistema. Validación de los lineamientos de OWASP y principalmente en el conjunto de 10 vulnerabilidades que debe resistir como mínimo en un sistema de información.
Performance		Tiempos de respuesta	Uso de librerías que optimicen el performance en acceso a datos. Utilización de comunicación por medio de JSON desde la capa más baja (datos) hasta la capa más alta (UI). Uso de archivos en formato Parquet para mejorar la velocidad de análisis basado en columnas para Hadoop (si la información se presenta en muchas columnas)
Mantenibilidad		DevOps	Continuidad en el uso del VSTS para todo el código generado

4.6. ARQUITECTURA ACTUAL Y DE REFERENCIA

ESTADO ACTUAL (AS IS)

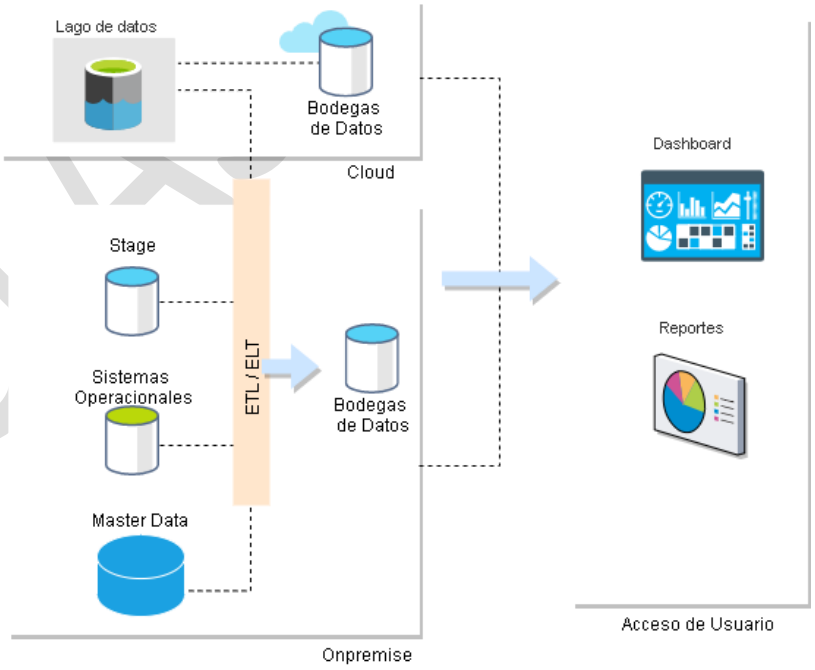
CLASIFICADA



La arquitectura de datos actual es de tipo hibrida, las bases de datos operacionales, bases de datos intermedias, maestra de datos y bodegas están en ambientes Microsoft On-premise, para analítica avanzada se tiene implementado un lago de datos en Azure Cloud, para la visualización se tiene la plataforma PowerBI.

- **ARQUITECTURA DE REFERENCIA**

ESTADO FUTURO (TO BE)



Cualquier copia impresa de este documento se considera como COPIA NO CONTROLADA.
LOS DATOS PROPORCIONADOS SERÁN TRATADOS DE ACUERDO A LA POLÌTICA DE TRATAMIENTO DE DATOS PERSONALES DEL ICBF Y A LA LEY 1581 DE 2012

La implementación de cualquier proyecto de Analítica debe tener en cuenta la arquitectura actual, las tácticas de arquitectura, la integración y la interoperabilidad, así como, la compatibilidad de las plataformas actuales de datos. La analítica avanzada y la creación de indicadores, deben basarse en determinar las fuentes de datos de los sistemas operaciones, bases de datos intermedias, bodegas de datos, lago de datos, debe contar con un modelo de integración basado en procesos de extracción, transformación y carga, también es recomendable realizar la integración mediante servicios de negocio usando Web Services SOAP, y/o a través de APIs de integración (comúnmente utilizando HTTP), que permitirán el almacenamiento de estos datos a trabajar. Por último, la visualización de información debe tener en cuenta la segmentación de dicha información (según los requerimientos funcionales del ICBF y las tácticas de usabilidad descritas en este documento).

5. CONTROL DE CAMBIOS

Fecha	Versión	Descripción del Cambio
N/A	N/A	N/A

CLASIFICADA