INFORME DE ACTIVIDADES

Plataforma de Analítica de Datos del Observatorio de Salud

Dirección Territorial de Salud de Caldas

Centro de Bioinformática y Biología Computacional de Colombia – BIOS

Septiembre de 2026

1. Información general

Proyecto / convenio	Convenio DTSC – BIOS: fortalecimiento y modernización del Observatorio mediante una plataforma para la gestión, procesamiento, análisis y visualización de datos.
Entidad beneficiaria	Dirección Territorial de Salud de Caldas
Entidad ejecutora	Centro de Bioinformática y Biología Computacional de Colombia – BIOS
Período reportado	21 de agosto al 09 de septiembre
Fecha estimada de inicio del proyecto	21 de agosto de 2026

2. Resumen ejecutivo

Durante las primeras semanas de ejecución se desarrollaron los instrumentos técnicos y de gestión requeridos para organizar el inicio del proyecto, precisar el alcance metodológico y establecer las bases para el tratamiento, gobierno y seguridad de la información. El trabajo comprendió la planeación detallada de actividades (Cronograma de trabajo), la definición de la metodología de ejecución, la estructuración del modelo de gobernanza de datos y la realización de una sesión formal de levantamiento de requerimientos con la Dirección Territorial de Salud de Caldas.

Como resultado, se consolidó una ruta de trabajo articulada en etapas, componentes y productos; se definió un ciclo operativo de ejecución y validación; se establecieron responsabilidades y criterios para la gestión de fuentes, calidad, linaje, metadatos, acceso y privacidad; y se documentaron los requerimientos funcionales y técnicos preliminares que orientarán el diseño y desarrollo de la Plataforma de Analítica de Datos del Observatorio de Salud.

Dichas actividades se consideran ejecutadas en cuanto a la elaboración de los productos documentales previstos para esta fase. Los asuntos que permanecen por definir como la priorización definitiva de fuentes, reglas de calidad específicas, indicadores, variables objetivo y capas geográficas corresponden a insumos y decisiones que deben continuar precisándose durante las siguientes etapas del proyecto, conforme se reciban y caractericen las fuentes de información.

3. Objetivo del informe

Presentar de manera consolidada le informe de las actividades desarrolladas por BIOS para dar cumplimiento a la fase inicial de diseño, estructuración y seguridad del proyecto, describiendo el propósito, contenido, resultados y evidencia de los documentos elaborados, así como su articulación con las siguientes etapas de ejecución.

4. Actividades desarrolladas

4.1 Elaboración del Cronograma de Trabajo Detallado

Se estructuró el Cronograma Institucional de Ejecución como instrumento de planeación y seguimiento del proyecto. El documento organiza la ejecución en fases y períodos, identifica componentes, entregables, fechas estimadas, responsables, dependencias, productos o hitos esperados, pagos asociados, porcentaje de avance y estado.

Para la fase inicial se definieron actividades relacionadas con la reunión de inicio, un moedelo de gobernanza de los datos, la estructuración metodológica, el levantamiento de requerimientos, la socialización de los productos y su revisión y aprobación. El cronograma general proyecta el desarrollo, integración, validación, despliegue y transferencia de conocimiento.

4.2 Elaboración de la Metodología de Trabajo

Se elaboró la Metodología de Trabajo para convertir el alcance técnico del proyecto en una ruta de ejecución clara, verificable y trazable. El documento adopta un enfoque ágil, modular y progresivo, con validaciones sucesivas y coordinación continua entre BIOS y la Dirección Territorial de Salud de Caldas.

La metodología establece un ciclo operativo de seis momentos: alistar, diseñar, construir, validar, desplegar y transferir/operar. Este ciclo se aplica de manera iterativa a los tres componentes tecnológicos de la solución y permite incorporar la validación desde el levantamiento inicial, en lugar de concentrarla únicamente al cierre.

1. ALISTAR Requerimientos, fuentes, accesos y criterios de trabajo.	**2. DISEÑAR** Arquitectura, reglas, modelos de datos, flujos y prototipos.	**3. CONSTRUIR** Pipelines, módulos analíticos, aplicación/web y prototipos.
4. VALIDAR Demostraciones, QA, revisión funcional y observaciones.	**5. DESPLEGAR** Configuración, documentación y entrega técnica.	**6. TRANSFERIR Y OPERAR** Capacitación, alojamiento soberano y soporte posterior.

Adicionalmente, se definieron los principios de modularidad, iteración y validación progresiva, trazabilidad, calidad y gobierno del dato, seguridad y soberanía, interoperabilidad estructurada, usabilidad y decisión basada en evidencia. La metodología desarrolla el flujo de datos en zonas Bronze, Silver y Gold del Data Lakehouse y describe la ruta de construcción y validación de los módulos de calidad estadística y gobierno, vigilancia epidemiológica espacial, AutoML, insights con IA local, captura web y móvil, y modernización de vistas del Observatorio Social.

4.3 Estructuración del Modelo de Gobernanza de Datos, Seguridad y Anonimización

Se estructuró un Modelo de Gobernanza de Datos orientado a establecer principios, responsabilidades, autoridades y mecanismos de decisión para la administración de los datos del Observatorio durante su ciclo de vida. El modelo plantea que la gobernanza comprende tanto la gestión funcional como la tecnológica, e incluye criterios de calidad, clasificación, acceso, compartición, metadatos, trazabilidad y uso autorizado.

El modelo define cuatro roles principales: propietario del dato, responsable funcional, responsable técnico o custodio y usuarios de datos. Esta asignación permite diferenciar quién toma decisiones sobre las fuentes, quién valida su significado y reglas de calidad, quién administra los mecanismos tecnológicos y quién utiliza la información bajo las autorizaciones correspondientes.

Eje	Contenido desarrollado	Resultado esperado
Roles y responsabilidades	Propietario del dato, responsable funcional, custodio técnico y usuarios.	Decisiones y responsabilidades claramente asignadas.

Eje	Contenido desarrollado	Resultado esperado
Caracterización de fuentes	Origen, naturaleza, ubicación, temporalidad, finalidad, proceso y productos derivados.	Contexto suficiente para clasificar, utilizar y controlar cada fuente.
Calidad de datos	Completitud, validez, consistencia, unicidad, exactitud, oportunidad, integridad y trazabilidad.	Reglas de calidad con diferenciación entre decisión funcional e implementación técnica.
Linaje y metadatos	Fuente, fecha, transformaciones, versión, producto derivado y dependencias.	Trazabilidad desde el origen hasta los productos de información.
Seguridad y privacidad	Protección proporcional a sensibilidad, criticidad y riesgo; mínimo privilegio, necesidad de conocer, responsabilidad individual y trazabilidad.	Acceso controlado y uso de la información conforme a autorizaciones.
Anonimización	Tratamiento de fuentes sensibles condicionado a su estructura, sensibilidad y restricciones legales, contractuales o institucionales.	Aplicación de medidas de anonimización cuando corresponda y ajuste conforme se caractericen las fuentes.

La metodología complementa este modelo al señalar que la definición concreta de reglas de seguridad y anonimización depende de las bases efectivamente suministradas, su estructura, sensibilidad y condiciones de uso. Por esta razón, el marco se establece en la fase inicial y se ajusta posteriormente conforme se caractericen las fuentes.

4.4 Levantamiento y documentación de requerimientos

El 25 de agosto de 2026 se realizó una sesión virtual de levantamiento de requerimientos con participación de representantes de la Dirección Territorial de Salud de Caldas y BIOS. La sesión permitió documentar el contexto inicial del Observatorio, precisar necesidades funcionales y técnicas y establecer asuntos que requieren definición conjunta durante la ejecución.

Durante la sesión se informó que el Observatorio administra aproximadamente 60 fuentes de información, internas y externas. Se identificaron como relevantes fuentes como DANE y SISPRO, y se reconoció la existencia de bases con restricciones especiales relacionadas, entre otros asuntos, con víctimas, menores de edad, información del ICBF y datos epidemiológicos sensibles. Esta condición condujo a incorporar explícitamente requisitos de confidencialidad, protección y anonimización.

Grupo de requerimientos	Aspectos principales identificados

Grupo de requerimientos	Aspectos principales identificados
Inventario y priorización de fuentes	Inventario de fuentes asociadas a tableros; priorización por importancia epidemiológica y territorial, disponibilidad, calidad, periodicidad, estructura, acceso y restricciones.
Confidencialidad y anonimización	Anonimización de fuentes sensibles cuando lo exijan restricciones legales, contractuales o institucionales; revisión particular de RUV/UARIV, ICBF, menores, víctimas e información epidemiológica sensible.
Integración y duplicidad	Admisión de estructuras y variables diferentes; reglas de identificación, cruce, homologación y tratamiento de registros relacionados.
Calidad de datos	Trazabilidad del estado de validación; reglas de tipo, valores inválidos, vacíos, duplicados, rangos, formatos y consistencia; gestión de inconsistencias y alertas.
Roles y permisos	Administrador, usuario de cargue/actualización y usuario de consulta; acceso condicionado por sensibilidad y funciones.
Alertas de actualización	Registro de periodicidad, última actualización y próxima fecha prevista de corte, con alertas anticipadas.
Analítica e IA	Generación de insights y resúmenes con IA generativa; definición de eventos, variables objetivo, indicadores, métricas, tableros y alertas analíticas.
Georreferenciación	Identificación y priorización de capas geográficas para visualización y análisis territorial.
Documentación y trazabilidad	Relación documentada entre fuentes, variables, procesos, indicadores, modelos y productos de visualización.

5. Productos y evidencias generadas

Producto	Finalidad	Principales contenidos	Estado

Producto	Finalidad	Principales contenidos	Estado
Cronograma Institucional de Ejecución	Organizar temporalmente la ejecución y seguimiento.	Fases, actividades, fechas, responsables, dependencias, hitos, pagos y estados.	**Elaborado**
Metodología de Trabajo	Definir la ruta metodológica para diseño, desarrollo, validación, despliegue y operación.	Enfoque ágil y modular, ciclo operativo, componentes A/B/C, QA, seguridad, riesgos y trazabilidad.	**Elaborado**
Modelo de Gobernanza de Datos	Establecer el marco de gestión, responsabilidad y control de datos.	Roles, fuentes, calidad, activos, linaje, metadatos, seguridad, privacidad y acceso.	**Elaborado**
Acta de Levantamiento de Requerimientos	Documentar el entendimiento inicial y los requerimientos del proyecto.	Contexto, alcance preliminar, requerimientos, criterios de aceptación y pendientes por definir.	**Elaborada**

6. Conclusiones

La fase inicial permitió establecer una base común de planeación, coordinación, metodología y gobierno de datos para orientar el desarrollo de la Plataforma de Analítica de Datos del Observatorio de Salud.

Los documentos elaborados se encuentran articulados entre sí: el cronograma define cuándo y bajo qué dependencias se ejecutan las actividades; la metodología establece cómo se desarrollarán; el modelo de gobernanza determina responsabilidades y controles sobre la información; y el acta de requerimientos documenta qué necesidades y condiciones deben atenderse.

El trabajo realizado incorpora desde el inicio criterios de trazabilidad, calidad, seguridad, confidencialidad, privacidad y tratamiento de información sensible, reconociendo que las reglas específicas de anonimización deberán ajustarse conforme se caractericen las bases de datos efectivamente suministradas.

Con la elaboración de los cuatro productos documentales, se cuenta con la evidencia necesaria para acreditar el desarrollo de la actividad reportada y avanzar a las etapas de calidad e ingeniería de ingesta, desarrollo, analítica, pruebas, despliegue y transferencia de conocimiento.

Recibido: 09-09-2026

Alonso Jiménez Castañeda
CC75065451