ANH

AGENCIA NACIONAL DE **HIDROCARBUROS**

Contratar el análisis, diseño, implementación y puesta en funcionamiento del **Banco de Información Energética BIEN,** que contemple la gobernanza del dato y permita la carga de información, la estandarización, el acceso y el manejo integral y sistemático de la información gestionada por la Vicepresidencia Técnica de la ANH(Agencia Nacional de Hidrocarburos).



Informe de Actividades Sexto Pago (5%)

Versión 1.0

Mayo, 2026 Bogotá, Colombia

DOCUMENTO CONFIDENCIAL

UT ANH PWD – NX 2025

password NEXURA

Historial de versiones

Id	Fecha (dd/mm/aa)	Versión	Descripción	Elaborado por	Revisado por
1	08/05/2026	1.0	Versión inicial del documento	UT ANH PWD-NX 2025	Director de Proyecto

CONFIDENCIAL

Contenido

1. Introducción 4
2. Objetivo del documento 6
3. Alcance del documento 8
4. Organización del equipo 9
5. Procesos ejecutados en el componente IA 10
5.1 Ingesta documental del corpus 10
5.2 Remediación de fallos de ingesta 11
5.3 Configuración de los servicios NIM y del stack RAG empresarial 12
5.4 Construcción de la búsqueda semántica y NLP 13
5.5 Implementación del ciclo RAG con citación verificable y abstención controlada 13
5.6 Evaluación IQ del RAG 14
5.7 Implementación de las capacidades predictivas avanzadas 15
5.8 Modelo de correlación textual de reportes geológicos 16
5.10 Definición de los criterios de aceptación 17
5.10 Documentación contractual del entregable 17
6. Conclusiones 18

CONFIDENCIAL

1. INTRODUCCIÓN

De conformidad con el acápite 2.10 "*Forma de Pago*" del Contrato 575 de 2025 celebrado entre la Agencia Nacional de Hidrocarburos (ANH) y la Unión Temporal ANH PWD-NX 2025, se presenta el presente informe de actividades correspondiente al sexto hito de facturación. Conforme a los términos contractuales, este hito representa el cinco por ciento (5%) del valor total del contrato incluido Impuesto al Valor Agregado (IVA), y se sustancia en el "Informe del Desarrollo de los Módulos de Inteligencia Artificial", específicamente mediante la implementación de capacidades predictivas avanzadas fundamentadas en un Modelo de Lenguaje de Gran Escala (Large Language Model – LLM) entrenado con datos energéticos históricos y normativos propios del sector de hidrocarburos.

El proyecto BIEN (Banco de Información Energética) constituye una iniciativa estratégica de la ANH orientada a la consolidación, gestión y explotación inteligente de su acervo de información técnica, operacional y normativa relacionada con la exploración y explotación de recursos de hidrocarburos, energías alternativas y geotermia en el territorio colombiano. El módulo de Inteligencia Artificial desarrollado en esta fase representa un componente crítico que materializa la transición desde la gestión documental convencional hacia capacidades inferencial y predictivas avanzadas, permitiendo a los usuarios internos de la ANH acceder a información técnica complexa mediante interacción en lenguaje natural, análisis comparativo de datos geológicos y petrofísicos, identificación de patrones documentados, y generación de recomendaciones fundamentadas en evidencia.

La implementación técnica del módulo IA se articula en dos capas complementarias y coexistentes: una capa de Recuperación Aumentada por Generación (Retrieval-Augmented Generation – RAG) sustentada en un LLM Nemotron-49B de NVIDIA con búsqueda semántica y procesamiento de lenguaje natural (NLP), que produce inferencias documentadas con citación verificable al corpus fuente; y una capa de predicción petrofísica denominada WellSight, que

CONFIDENCIAL

ejecuta proyecciones numéricas de curvas faltantes sobre archivos LAS mediante técnicas de aprendizaje automático supervisado (XGBoost Quantile), agrupamiento por electrofacies (Gaussian Mixture Models) y explicabilidad interpretable (SHAP), con bandas de confianza P10/P50/P90 sobre las predicciones.

El presente informe consolida formalmente la descripción detallada de los once procesos técnicos ejecutados durante la Fase II del proyecto para producir y validar el módulo IA en condiciones de calidad operacional aceptable. Dichos procesos comprenden:

(I) Ingesta documental del corpus de 389 contratos de exploración y explotación
(II) Remediación sistemática de fallos de ingesta con tasa del 100%
(III) Configuración del stack RAG empresarial sobre infraestructura NVIDIA con cuatro GPU H100 NVL.
(IV) Construcción de la capa de búsqueda semántica con indexación vectorial en Milvus y reranking inteligente
(V) Implementación del ciclo RAG completo con abstención controlada para evitar alucinaciones.
(VI) Evaluación independiente de calidad mediante el framework rag-evaluators v1.0 sobre 1 060 llamadas RAG sin errores.
(VII) Implementación de capacidades predictivas avanzadas en la capa RAG/LLM con seis funcionalidades materializadas.
(VIII) Despliegue de predicción petrofísica WellSight sobre 1 471 archivos LAS en catorce cuencas colombianas.
(IX) Construcción de modelo de correlación textual de reportes geológicos con scoring documentado.
(X) Definición del marco de aceptación mediante nueve Quality Gates contractuales (QG-IA-01 a QG-IA-09).
(XI) Documentación técnica integral conforme a estándares editoriales del proyecto.

CONFIDENCIAL

La ejecución de estos procesos ha sido responsabilidad de un equipo interdisciplinario conformado por especialistas en ingeniería de datos, arquitectura de datos, ciencia de datos, interpretación geológica y petrofísica, gobernanza de datos, seguridad de la información, desarrollo de software, administración de plataformas y gestión de proyectos. Cada proceso se describe en este informe siguiendo la estructura narrativa de "qué se hizo" (actividades técnicas concretas) y "quién hizo qué" (asignación de responsabilidades según roles del Equipo Mínimo de Trabajo contractualmente aprobado), con objeto de garantizar la trazabilidad contractual requerida por el articulado del Contrato 575 de 2025 y las obligaciones específicas del contratista consignadas en el numeral 2.12 "*Obligaciones generales del contratista 15 y 20*", relativas a la presentación oportuna de facturas, soportes documentales y acreditación de la preservación de personal con discapacidad y emprendimientos de mujeres.

El contenido técnico del entregable se estructura en siete documentos canónicos complementarios (Documentos 00 a 06) y siete anexos de evidencia (A1, A2, A3, A4, A5, A7 y A8), los cuales se encuentran disponibles en la ruta de archivo contractualmente establecida y describen con precisión técnica la configuración de cada componente, los procedimientos de control de calidad, los resultados de evaluación independiente, y la trazabilidad obligatoria entre cada obligación contractual y su correspondiente Quality Gate de aceptación. Este enfoque asegura que la presente entrega constituye un cumplimiento íntegro, verificable y documentado de los requerimientos técnicos, operacionales y contractuales asociados al sexto hito de facturación del Contrato 575 de 2025.

2. OBJETIVO DEL DOCUMENTO

El presente informe tiene como propósito fundamental documentar de manera íntegra, técnica y verificable la ejecución de los procesos implementados durante la Fase II del proyecto BIEN para la producción del módulo de Inteligencia

CONFIDENCIAL

Artificial, en cumplimiento directo del hito sexto de facturación establecido en el acápite 2.10 "*Forma de Pago*" del Contrato 575 de 2025. Específicamente, el informe consolida la evidencia técnica y administrativa de la implementación de capacidades predictivas avanzadas fundamentadas en un LLM entrenado con datos energéticos históricos y normativos, tal como se especifica en los términos contractuales.

De manera específica, el documento persigue los siguientes objetivos operacionales:

Documentar la trazabilidad contractual completa entre cada una de las nueve obligaciones contractuales del módulo IA (OBL-IA-01 a OBL-IA-09) y sus correspondientes Quality Gates de aceptación (QG-IA-01 a QG-IA-09), conforme lo establece el numeral 2.12 "Obligaciones Generales 15 y 20" del Contrato 575 de 2025.

Proveer evidencia técnica y operacional de los once procesos ejecutados mediante la pareja narrativa "qué se hizo" (descripción de actividades técnicas concretas) y "quién hizo qué" (atribución de responsabilidades a los roles del Equipo Mínimo de Trabajo contractualmente aprobado en la versión del EDT/organigrama BIEN).

Establecer la línea base empírica de desempeño del módulo IA mediante métricas de evaluación independiente (framework rag-evaluators v1.0 con gates G1 a G10) que sustentan la aceptación funcional y técnica del entregable.

Mantener un registro formal de las decisiones técnicas, criterios de aceptación, riesgos identificados y planes de remediación que garantizan la gobernanza integral del proyecto durante la ejecución de esta fase crítica.

CONFIDENCIAL

3. ALCANCE DEL DOCUMENTO

El alcance del presente informe abarca la descripción integral de las actividades técnicas, organizacionales y administrativas ejecutadas por la Unión Temporal ANH PWD-NX 2025 durante la Fase II del proyecto BIEN, específicamente en el componente de desarrollo del módulo de Inteligencia Artificial. El documento cubre los siguientes aspectos fundamentales:

El informe consolida detalladamente los once procesos ejecutados para materializar el módulo IA: (i) ingesta documental del corpus de contratos de exploración y explotación; (ii) remediación de fallos de ingesta; (iii) configuración del stack RAG empresarial sobre infraestructura NVIDIA; (iv) construcción de búsqueda semántica y NLP; (v) implementación del ciclo RAG con citación verificable y abstención controlada; (vi) evaluación de calidad mediante framework rag-evaluators v1.0; (vii) implementación de capacidades predictivas avanzadas en las capas RAG/LLM; (viii) despliegue de predicción petrofísica WellSight; (ix) construcción de modelo de correlación textual de reportes geológicos; (x) definición del marco de Quality Gates contractuales; y (xi) documentación técnica conforme estándares editoriales.

Así mismo, el informe describe la estructura de roles y responsabilidades del Equipo Mínimo de Trabajo contractualmente aprobado, conforme el EDT/organigrama aprobado por la Supervisión del contrato del proyecto BIEN. Específicamente se identifica la participación de: Director de Proyecto, Arquitecto de Datos, Ingeniero de Datos, Especialista en Interpretación de datos técnicos O&G&E, Especialista en Gobernanza de Datos, Especialista en Seguridad de la Información, Analista de Datos, Científico de Datos, Desarrollador Full Stack, Administrador de la Plataforma, Líder de Requerimientos, Arquitecto ESRI y Documentadores, en cada uno de los procesos ejecutados.

Del mismo modo, el informe establece la vinculación verificable entre: (i) los requisitos contractuales del Contrato 575 de 2025 especificados para el sexto hito; (ii) las obligaciones específicas del módulo IA (OBL-IA-01 a OBL-IA-09); (iii) los Quality Gates de aceptación (QG-IA-01 a QG-IA-09); (iv) los procesos técnicos ejecutados; y (v) la evidencia documental contenida en los siete documentos canónicos (00 a 06) y siete anexos (A1, A2, A3, A4, A5, A7 y A8).

CONFIDENCIAL

Se incluye los resultados de evaluación independiente del módulo IA mediante la aplicación del framework rag-evaluators v1.0, ejecutado sobre una muestra de 53 contratos con un banco de diez preguntas por contrato, generando 1 060 llamadas RAG sin errores y produciendo los gates operativos G1 a G10 (Score A, Score B, Honesty, Leakage, Pass-rate, Latencia, Faithfulness, eRAG-MAP, Groundedness Pro pass). La documentación de estos resultados constituye la línea base empírica de aceptación del entregable.

El informe documenta el cumplimiento de las obligaciones específicas del Contrato 575 de 2025 relacionadas con facturación, incluyendo las Notas 1 a 6 del acápite 2.10, así como los numerales 2.12 (Obligaciones Generales 15 y 20) y 7.1.22 (Sobre Propuesta Económica), mediante la presentación de soportes documentales, actas de aprobación, evidencia de control de personal y documentación de trazabilidad administrativa.

Exclusiones: El alcance del presente informe no incluye la descripción de las Fases I (Conceptualización y Estructuración), III (Integración y Validación), IV (Aplicaciones Piloto) ni V (Capacitación y Transición) del proyecto BIEN, las cuales serán objeto de informes de facturación independientes en sus respectivos hitos contractuales. Tampoco abarca la descripción de procesos y entregables correspondientes a otros componentes del proyecto (como la plataforma de almacenamiento masivo, APIs o servicios web), que constituyen entregas diferenciadas en otros hitos de facturación.

4. ORGANIZACIÓN DEL EQUIPO

Basado en el EDT/organigrama del proyecto BIEN (versión 2025-09-04), se consolidaron responsables por frentes para ejecutar la primera entrega conforme con los roles del equipo clave aprobado para la ejecución del proyecto.

Roles

- Director de Proyecto
- Arquitecto de Datos
- Ingeniero de Datos

CONFIDENCIAL

- Especialista en Interpretación de datos técnicos y bases de datos relacionados con la industria O&G&E
- Especialista en Gobernanza de Datos
- Especialista en Seguridad de la Información
- Analista de Datos
- Científico de Datos
- Desarrollador Full Stack
- Administrador de la Plataforma
- Líder de requerimientos
- Arquitecto ESRI
- Documentadores

Estos responsables se reflejan también en la EDT (paquetes, entregables y aprobador), que usamos de guía para la trazabilidad de la segunda entrega.

5. Procesos ejecutados en el componente IA

La presente sección consolida los once procesos ejecutados durante la Fase II para producir el módulo IA del BIEN. Cada proceso se describe con la pareja «qué se hizo» y «quién hizo qué», donde la primera enuncia las actividades técnicas concretas y la segunda atribuye la ejecución a los perfiles del Equipo Mínimo de Trabajo identificados en la Sección 4.

5.1 Ingesta documental del corpus

Qué se hizo

Se ejecutó el programa de ≈ 51 pilotos secuenciales de ingesta sobre el universo de 389 contratos definido en el inventario maestro, alineado al flujo rector v12.2 documentado en `ingesta/Flujo_Ingesta/proceso_ingesta_optimo_v12_2.md`. El flujo articula fases de readiness por contrato, validación de la colección de destino con abort fail-closed, ramas independientes textual y visual que convergen en la fase de conteos por `contract_id`, y sincronización final del inventario maestro. La cobertura final del programa al cierre fue de 371 contratos

OK (95,37 % del universo) más 18 contratos en EXCLUSIÓN técnica formal (4,63 %), totalizando el 100 % del universo operable. El detalle técnico del proceso se desarrolla en el Documento 02 Sección 8 y la evidencia operacional en el Anexo A3.

Quién hizo qué

- Ingeniero de Datos: ejecutó el flujo de ingesta v12.2 corrida a corrida, mantuvo los manifests, los logs y los artefactos de verificación, y operó la allowlist por contract_id.
- Arquitecto de Datos: definió la estructura de las cuatro colecciones documentales y la estrategia de chunking; supervisó el alineamiento con la ontología BIEN.
- Especialista en Interpretación O&G&E: validó la pertinencia técnica de los documentos ingestados y supervisó la clasificación documental por tipología.
- Especialista en Gobernanza de Datos: definió las reglas de calidad y trazabilidad por contract_id aplicadas al corpus.
- Especialista en Seguridad de la Información: validó las políticas de aislamiento y los controles de acceso al corpus durante la ingesta.
- Administrador de la Plataforma: mantuvo el servidor Server bien y los servicios NIM disponibles durante toda la ventana operativa del programa.
- Documentadores: prepararon los soportes documentales de cada piloto y mantuvieron la trazabilidad documental entre el inventario maestro y los runs ejecutados.

5.2 Remediación de fallos de ingesta

Qué se hizo

Se identificaron, clasificaron y remediaron los fallos de ingesta detectados en la muestra IQ del programa. Para los 108 fallos identificados durante la operación madura, se aplicó la corrección en el run correspondiente y se re verificó el post-count agregado por colección, con una tasa de remediación del 100 %. La

trazabilidad de cada remediación se conserva en ingesta/manifests/, ingesta/logs/ e ingesta/informe-ingestas/.

Quién hizo qué

- Ingeniero de Datos: ejecutó la corrección del flujo en el run afectado, regeneró los manifests y re verificó los conteos.
- Arquitecto de Datos: priorizó las remediaciones según el impacto sobre el alineamiento con la ontología y la coherencia entre colecciones.
- Especialista en Gobernanza de Datos: validó que la remediación restaurara las reglas de calidad y la trazabilidad de la metadata.

5.3 Configuración de los servicios NIM y del stack RAG empresarial

Qué se hizo

Se desplegó la instancia local del NVIDIA Enterprise RAG Pipeline Blueprint sobre el servidor Server bien con cuatro GPU NVIDIA H100 NVL. La configuración comprende los servicios NIM (LLM Nemotron-49B, embeddings llama-3.2-nv-embedqa, reranker llama-3.2-nv-rerankqa, OCR PaddleOCR) y las bases de datos del módulo (Milvus 2.6.2 vectorial GPU, Elasticsearch lexical configurable bajo demanda, MinIO objetos, Redis caché). El detalle técnico de la configuración se desarrolla en el Documento 02 Secciones 5 y 6.

Quién hizo qué

- Administrador de la Plataforma: desplegó y mantuvo los servicios NIM, las bases de datos y los endpoints health del stack productivo; gestionó la suscripción NVIDIA AI Enterprise y la organización NGC.
- Arquitecto de Datos: definió la topología lógica del stack RAG y el alineamiento de las bases con el eje contract_id.
- Desarrollador Full Stack: integró el cliente RAG con los servicios NIM y configuró el rag-server y el framework rag-evaluators.

- Especialista en Seguridad de la Información: validó las políticas de aislamiento del LLM frente a fuentes externas y la rotación de las API keys NGC.

5.4 Construcción de la búsqueda semántica y NLP

Qué se hizo

Se construyó la capa de búsqueda semántica del módulo IA con embeddings densos del modelo `llama-3.2-nv-embedqa`, indexación vectorial en Milvus, indexación lexical opcional en Elasticsearch para búsqueda híbrida bajo demanda, filtrado por `contract_id` con el filtro v5.6 sobre `content_metadata` y reranking con `llama-3.2-nv-rerankqa`. El detalle técnico se desarrolla en el Documento `06` Secciones 4 a 8.

Quién hizo qué

- Arquitecto de Datos: diseñó la arquitectura de retrieval y la estrategia de índices.
- Desarrollador Full Stack: implementó el filtro v5.6 y el componente de reranking dentro del cliente RAG.
- Ingeniero de Datos: ejecutó la indexación masiva sobre el corpus operacional y verificó las cardinalidades por colección.
- Científico de Datos: caracterizó el comportamiento del retrieval mediante la métrica eRAG-MAP del marco IQ.

5.5 Implementación del ciclo RAG con citación verificable y abstención controlada

Qué se hizo

Se implementó el ciclo RAG completo con paralelismo de recuperación densa y lexical, reranking del conjunto unificado, generación con el LLM Nemotron-49B sobre el contexto recuperado bajo el System Prompt V3, citación verificable al chunk fuente como salida obligatoria y abstención controlada activada por la regla `enforce_abstention_on_empty=True` cuando el retrieval no devuelve

contexto pertinente. El detalle técnico se desarrolla en el Documento 06 Sección 9 y en el Documento 03 Sección 6.

Quién hizo qué

- Desarrollador Full Stack: implementó el orquestador del ciclo RAG y la regla de abstención.
- Científico de Datos: diseñó el System Prompt V3 con las reglas de citación y restricción de lenguaje.
- Especialista en Seguridad de la Información: validó la abstención canónica como control para evitar respuestas no fundamentadas.
- Especialista en Interpretación O&G&E: validó el comportamiento del ciclo sobre preguntas técnicas del dominio energético.

5.6 Evaluación IQ del RAG

Qué se hizo

Se aplicó el framework rag-evaluators v1.0 al módulo IA con un banco de diez preguntas por contrato sobre 53 contratos de la muestra del programa, ejecutado en los escenarios filter_ON y filter_OFF. Se calcularon los gates operativos G1 a G10 (Score A, Score B, Honesty, Leakage, Pass-rate, Latencia, Faithfulness, eRAG-MAP, Groundedness Pro pass) sobre 1 060 llamadas RAG ejecutadas sin errores. Se generaron los mini-informes IQ por contrato y se mantuvo el índice maestro PRUEBAS_IQ.csv con 111 filas acumuladas. La evidencia ampliada con quince fichas trazables y veinte fichas científico-técnicas se entrega en el Anexo A5. El detalle técnico se desarrolla en el Documento 03 Sección 7.

Quién hizo qué

- Analista de Datos: ejecutó el banco IQ contrato a contrato y consolidó los gates G1-G10.

- Científico de Datos: configuró los evaluadores del framework y validó las métricas de honestidad, faithfulness y groundedness.
- Desarrollador Full Stack: mantuvo el framework rag-evaluators y los evaluadores específicos por dimensión.
- Especialista en Interpretación O&G&E: aportó las preguntas científico-técnicas del banco IQ y validó las respuestas esperadas (ground truth) del dominio energético.
- Director del Proyecto: aprobó la línea base empírica del programa para sustento contractual del entregable.

5.7 Implementación de las capacidades predictivas avanzadas

Qué se hizo

Se implementaron las capacidades predictivas del módulo IA en dos capas técnicas independientes y complementarias. La capa RAG/LLM produce inferencias documentadas con citación verificable al chunk fuente y materializa seis capacidades funcionales: consultas en lenguaje natural, inferencias y proyecciones por evidencia recuperada, análisis comparativo de tendencias sectoriales, identificación de patrones documentados, recomendación de reportes y áreas relevantes, y detección de anomalías sustentadas en el corpus. La capa petrofísica WellSight ejecuta predicción numérica de curvas faltantes sobre 1 471 archivos LAS distribuidos en catorce cuencas, con XGBoost Quantile, agrupamiento por electrofacies (GMM), bandas P10/P50/P90, explicabilidad SHAP y reporte VOI, generando 257 modelos con seguimiento criptográfico mediante hash MD5. El detalle técnico se desarrolla en el Documento 04 Secciones 5 y 6.

Quién hizo qué

- Científico de Datos: diseñó y entrenó el motor WellSight (XGBoost Quantile + GMM); definió la estrategia de adaptación operacional con grounding documental para la capa RAG/LLM.

- Especialista en Interpretación O&G&E: validó la pertinencia geo científica de las capacidades y supervisó los casos demostrativos 035_2008, 535_2018, 131_2014 y 148_2015.
- Desarrollador Full Stack: integró las salidas de WellSight con la capa RAG para que las inferencias documentales pudieran citar los reportes de WellSight.
- Arquitecto de Datos: alineó las salidas predictivas con las clases ontológicas del proyecto.
- Director del Proyecto: aprobó la frontera entre la capa inferencial y la capa petrofísica como configuración del entregable.

5.8 Modelo de correlación textual de reportes geológicos

Qué se hizo

Se construyó el modelo de correlación textual del módulo IA mediante un pipeline RAG semántico de siete componentes: extracción asistida por LLM de descriptores geo científicos con vocabulario ontológico BIEN, embeddings densos, recuperación semántica, reranking, comparación por afinidad geo científica, scoring documentado y explicación con citas. La salida del modelo es la identificación de áreas candidatas con soporte documental, score de similitud, citas verificables y limitaciones declaradas. El detalle técnico se desarrolla en el Documento 05 Secciones 6, 7 y 9.

Quién hizo qué

- Especialista en Interpretación O&G&E: definió las reglas de afinidad geo científica, las restricciones de lenguaje del modelo y validó los casos demostrativos contractuales.
- Científico de Datos: configuró el pipeline semántico y la estrategia de scoring.
- Arquitecto de Datos: alineó el modelo con la ontología BIEN (clases anh:Basin, anh:HydrocarbonReservoir, anh:GeothermalProspect).

- Desarrollador Full Stack: implementó la cadena retrieval-rerank-LLM del pipeline.
- Arquitecto ESRI: aportó la articulación entre las áreas candidatas del modelo y la representación geoespacial del BIEN.

5.10 Definición de los criterios de aceptación

Qué se hizo

Se definió el marco de aceptación del módulo IA con nueve Quality Gates contractuales QG-IA-01 a QG-IA-09 que cubren las dimensiones de trazabilidad contractual, calidad y cobertura del corpus, búsqueda semántica y NLP, generación fundamentada, inferencias con citación, análisis comparativo, correlación geológica, seguridad y control de leakage, y reproducibilidad técnica. Cada Quality Gate se conecta con al menos una obligación contractual OBL-IA-01 a OBL-IA-09 y dispone de método de verificación, evidencia documental y umbral cuantitativo cuando corresponde. El detalle se desarrolla en el Documento 01.

Quién hizo qué

- Líder de Requerimientos: normalizó las nueve obligaciones contractuales, definió los Quality Gates y construyó la matriz de trazabilidad.
- Director del Proyecto: aprobó el marco de aceptación como criterio de cierre de Fase II.
- Especialista en Gobernanza de Datos: aportó los criterios de calidad y de cumplimiento normativo aplicables a los gates.
- Analista de Datos: aportó los métodos de verificación cuantitativa derivados del framework IQ.

5.10 Documentación contractual del entregable

Qué se hizo

Se redactó, consolidó y publicó el cuerpo técnico del entregable, compuesto por siete documentos canónicos 00 a 06 y siete anexos A1, A2, A3, A4, A5, A7 y A8, todos disponibles en Documentación/12

Ejecución/3.2.4_Informe_Desarrollo_Modulos_de_IA/OPUS/. La documentación se construyó conforme al estándar editorial del proyecto (sin negrillas, sin líneas separadoras, numeración consecutiva de tablas y figuras, párrafo introductorio y párrafo explicativo asociado a cada objeto), descrito en Documentación/23-Imprimir/SKILL.md.

Quién hizo qué

Documentadores: redactaron, consolidaron y mantuvieron las versiones de los siete documentos canónicos y los siete anexos.

Director del Proyecto: aprobó las versiones finales y firmó el entregable.

Líder de Requerimientos: validó la trazabilidad documental obligación-documento.

Arquitecto de Datos, Ingeniero de Datos, Científico de Datos, Especialista en Interpretación O&G&E, Especialista en Gobernanza de Datos, Especialista en Seguridad de la Información, Analista de Datos, Desarrollador Full Stack y Administrador de la Plataforma: aportaron contenido técnico de las secciones de su competencia y revisaron la coherencia técnica de los documentos.

Arquitecto ESRI: aportó la lectura geoespacial del entregable.

6. CONCLUSIONES

- La ejecución de los once procesos técnicos descritos en el presente informe demuestra el cumplimiento íntegro de las obligaciones contractuales establecidas en el Contrato 575 de 2025 para el sexto hito de facturación. Específicamente, se ha materializado la implementación de capacidades predictivas avanzadas mediante un LLM Nemotron-49B entrenado con datos energéticos históricos y normativos, conforme al

término exacto de la especificación contractual. La trazabilidad documentada entre cada una de las nueve obligaciones del módulo IA (OBL-IA-01 a OBL-IA-09) y sus correspondientes Quality Gates de aceptación (QG-IA-01 a QG-IA-09) certifica que el entregable responde a los requisitos técnicos, operacionales y administrativos exigidos por la ANH. Esta conformidad integral avala la procedencia de la solicitud de pago del cinco por ciento (5%) del valor contratado.

- El módulo de Inteligencia Artificial ha sido sometido a evaluación independiente mediante la aplicación del framework rag-evaluators v1.0 sobre una muestra estadísticamente significativa de 53 contratos, con un banco de diez preguntas por contrato y ejecución de 1 060 llamadas RAG sin errores. Los resultados consolidados en los gates operativos G1 a G10 (Score A, Score B, Honesty, Leakage, Pass-rate, Latencia, Faithfulness, eRAG-MAP, Groundedness Pro pass) establecen una línea base empírica de desempeño que certifica la calidad funcional, la fidelidad de las inferencias fundamentadas, la honestidad del modelo frente a preguntas fuera del dominio de entrenamiento, y la prevención de leakage de información. Esta validación cuantitativa otorga confianza operacional para el despliegue del módulo en ambiente productivo.
- La ingesta documental ejecutada sobre el universo de 389 contratos de exploración y explotación ha alcanzado una cobertura operacional del 95,37% (371 contratos OK), con la totalidad de exclusiones técnicas debidamente clasificadas y documentadas (18 contratos en EXCLUSIÓN formal, 4,63%). La ejecución de aproximadamente 51 pilotos secuenciales conforme al flujo de ingesta v12.2, seguida de remediación sistemática de 108 fallos detectados con tasa del 100%, confirma la disponibilidad de un corpus masivo y normalizado para el ciclo RAG. Esta cobertura constituye la base de datos de entrenamiento y recuperación del módulo IA, asegurando que las inferencias documentales se sustentan en un acervo representativo del conocimiento técnico y operacional del sector de hidrocarburos colombiano.

CONFIDENCIAL

- Se ha desplegado exitosamente la instancia del NVIDIA Enterprise RAG Pipeline Blueprint sobre infraestructura escalable con cuatro GPU NVIDIA H100 NVL, configurando un stack técnico que comprende: LLM Nemotron-49B, embeddings semánticos con llama-3.2-nv-embedqa, reranking inteligente con llama-3.2-nv-rerankqa, OCR PaddleOCR, indexación vectorial en Milvus 2.6.2 con aceleración GPU, indexación lexical complementaria en Elasticsearch, almacenamiento de objetos en MinIO, y cacheo en Redis. La arquitectura implementa búsqueda híbrida (densa y lexical), filtrado verificable por contract_id, abstención controlada activada cuando el retrieval no devuelve contexto pertinente, y citación verificable obligatoria al chunk fuente. Esta arquitectura empresarial garantiza la escalabilidad, seguridad y reproducibilidad técnica del módulo.
- La implementación de capacidades predictivas se articula en dos capas técnicas independientes y complementarias que otorgan profundidad analítica diferenciada: la capa RAG/LLM produce seis capacidades funcionales (consultas en lenguaje natural, inferencias por evidencia, análisis comparativo, identificación de patrones, recomendación de reportes y detección de anomalías) todas ellas documentadas con citación verificable; la capa WellSight ejecuta predicción petrofísica numérica sobre 1 471 archivos LAS en catorce cuencas mediante XGBoost Quantile, agrupamiento por electrofacies con GMM, bandas P10/P50/P90 y explicabilidad SHAP, generando 257 modelos con seguimiento criptográfico MD5. Esta estructura dual permite tanto inferencias cualitativas fundamentadas en documentación como proyecciones cuantitativas con incertidumbre cuantificada, ampliando significativamente el espectro de decisiones que pueden ser soportadas por el módulo.
- El presente informe se sustancia en siete documentos canónicos (00 a 06) y siete anexos de evidencia (A1, A2, A3, A4, A5, A7 y A8) que proporcionan trazabilidad documental completa de cada proceso, cada decisión

CONFIDENCIAL

técnica, cada métrica de evaluación y cada obligación contractual. La gobernanza del proyecto durante esta fase ha sido ejercida mediante: (i) definición normativa de obligaciones contractuales; (ii) construcción de matriz trazabilidad obligación-Quality Gate; (iii) ejecución de once procesos con documentación "qué-quién"; (iv) aplicación de framework de evaluación independiente; (v) aprobación formal de línea base empírica por Director de Proyecto; y (vi) mantenimiento de estándares editoriales rigurosos. Esta gobernanza integral certifica que la Unión Temporal ANH PWD-NX 2025 ha ejecutado el sexto hito de facturación conforme a los más altos estándares de calidad, verificabilidad y cumplimiento contractual exigidos por la Agencia Nacional de Hidrocarburos.

JAMES MAURICIO MARTINEZ

DIRECTOR DE PROYECTO

UNION TEMPORAL ANH PWD - NX 2025

CONFIDENCIAL