

Contexto y Desafíos Iniciales en la Gestión del Dato

En un momento clave de transformación organizativa y avance tecnológico, la división E&P (Exploración y Producción) de CEPSA (MOEVE) se enfrentaba al reto de consolidar el conocimiento crítico sobre los datos, unificar procesos dispersos y garantizar una gestión fiable y coherente de la información que sustenta la toma de decisiones tanto operativas como estratégicas.

Para afrontar este desafío, CEPSA E&P puso en marcha un piloto de 12 semanas —que se completó en solo 8 semanas efectivas— con el objetivo de validar la capacidad de Anjana Data como plataforma central para el gobierno del dato en toda la organización.

Un Proyecto Impulsado por la Colaboración

El éxito y la rapidez con la que se ejecutó este proyecto —que logró completar en solo 8 semanas efectivas la definición, implantación, configuración y puesta en marcha de los casos de uso— fueron posibles gracias a una **colaboración ejemplar entre todos los equipos involucrados**. La implicación, la coordinación continua y el alto nivel de especialización técnica fueron determinantes para acelerar cada fase del piloto y, al mismo tiempo, evidencian que Anjana Data es una plataforma ágil, flexible y no especialmente compleja de adoptar cuando una organización decide impulsar de forma seria su iniciativa de Gobierno del Dato.

Participación de los equipos:

➤ Anjana Data

El equipo de Customer Success actuó como guía estratégico del piloto, liderando la definición del modelo y la configuración avanzada de la plataforma. A ello se sumó el apoyo del equipo de Global Support, que aportó el acompañamiento técnico necesario para garantizar una implantación fluida y sin fricciones.

➤ CEPSA E&P

La unidad IT-Digital impulsó el proyecto desde su origen, trabajando de manera coordinada con las áreas de negocio responsables de los casos de uso: los Activos de Producción en Norte de África y el equipo de Planning, Performance & Economics. Su conocimiento funcional y operativo fue clave para validar el gobierno end-to-end del dato.

➤ Mática Partners

Desde la Dirección de Gobierno del Dato, aportaron consultoría especializada tanto en la definición del modelo como en la ejecución operativa de la prueba de concepto, asegurando que las prácticas implantadas fueran sólidas, escalables y alineadas con estándares de gobernanza.

Caso de uso: Gobierno del Dato en Acción

Se ha implementado una solución sobre Anjana Data en modalidad SaaS híbrido, orientada a establecer un modelo de gobierno sólido y a ejecutar dos casos de uso productivos de alta complejidad, complementados con un análisis de viabilidad para incorporar capacidades avanzadas.

1. Profundización en el Modelo de Gobierno

Se ha definido y puesto en marcha un modelo de gobierno específico para Cepsa E&P, caracterizado por:

- **Roles y Responsabilidades:** Asignación clara de actores dentro de la plataforma, incluyendo Data Owners, Business Data Stewards, Data Consumers y DM Data Stewards, garantizando que cada activo cuente con un responsable identificado.
- **Workflows de Validación:** Implementación de flujos automatizados para la aprobación de activos. Cada dataset, término o KPI requiere validación por parte del Data Owner y del Data Steward antes de su publicación o modificación, asegurando calidad y confianza en el dato.
- **Metamodelo personalizado:** Configuración de entidades clave (Términos de negocio, Datasets, Procesos, Instancias, Soluciones, KPIs, DSAs, Reglas de Calidad...) con plantillas específicas para capturar metadatos técnicos, funcionales y de seguridad.

2. Implementación de Casos de Uso Críticos

Se ha logrado trazabilidad end-to-end en dos escenarios de alta complejidad técnica:

- **Caso de Uso 1: Extraction & Prospecting (Activos Norte de África)**

Gobierno del flujo del dato de producción “Daily Production Figures”, tanto operados por la compañía como por terceros, asegurando trazabilidad end-to-end en uno de los flujos más críticos para una empresa Oil & Gas. En el desarrollo del proyecto se ha trazado el dato desde su origen on-premise (Excel, SQL Server), pasando por las etapas de ingesta mediante scripts en Python y Power Automate, hasta su incorporación en las bases de datos del Data Lake en AWS (capas Landing y Staging) y su visualización final. Además, se ha documentado la lógica de transformación que convierte los datos crudos en reportes diarios de producción, aportando visibilidad completa sobre linaje, fuentes y responsables en cada etapa del ciclo.

- **Caso de Uso 2: Operational Consumption (Planning, Performance & Economics)**

Gobierno del ciclo de vida de datos financieros y de planificación. El flujo abarca desde la recolección en plantillas Excel y SharePoint (SPO), su carga automatizada en AWS S3, y su transformación en capas de negocio (PostgreSQL) para consumo en cuadros de mando Power BI. Anjana Data ha permitido visualizar dependencias entre Dataflows y orígenes subyacentes, reforzando la transparencia y la gobernanza.

3. Análisis de Viabilidad: Seguridad y Gestión Documental

Se ha realizado un análisis funcional para validar la capacidad de Anjana Data de integrar futuros alcances estratégicos:

- **Protección de Datos (Análisis DLP):** Se ha evaluado la viabilidad de incorporar el motor Microsoft Purview DLP, con el objetivo de que Anjana ingeste automáticamente las etiquetas de sensibilidad (Confidencial, Interno) y las refleje en el catálogo. Esto permitiría activar flujos de trabajo ante cambios de clasificación o detección de actividades sospechosas, reforzando la seguridad y el cumplimiento normativo.
- **Gestión Documental (Análisis Documentum D2):** Se ha evaluado la propuesta de integración con Documentum D2 para sincronizar metadatos de documentos no estructurados. El análisis confirma que Anjana puede actuar como punto único de acceso, permitiendo buscar documentos por sus metadatos funcionales sin necesidad de replicar.

Caso de uso: Gobierno del Dato en Acción

El proyecto se orientó a alcanzar y validar los siguientes objetivos clave:

- **Automatización en la captura de metadatos:** Implementación de plugins en la infraestructura AWS de CEPESA E&P para recolectar automáticamente información técnica de datasets y procesos, reduciendo la intervención manual y minimizando el riesgo de error humano.
- **Democratización del conocimiento:** Creación de un Glosario de Negocio y un catálogo de KPIs, que permite a los usuarios navegar de forma intuitiva entre conceptos funcionales y su correspondencia técnica, impulsando la transparencia y la comprensión del dato.
- **Garantía de propiedad del dato:** Asignación de Data Owners a todos los activos críticos catalogados, eliminando ambigüedades sobre la responsabilidad y asegurando una gestión clara y eficiente.

Arquitectura de la solución:

La solución se ha construido sobre una arquitectura SaaS híbrida, diseñada para garantizar la soberanía del dato y maximizar la automatización:

- **Núcleo de Gobierno (SaaS):** Instancia de Anjana Data desplegada en la nube, responsable de la gestión del metamodelo, los flujos de trabajo y el portal de usuarios.
- **Infraestructura de Plugins (AWS CEPESA E&P):** Servidores alojados en la red de CEPESA E&P que ejecutan los plugins de extracción. Estos componentes leen los metadatos de las fuentes (S3, SQL Server, Power BI, entre otras) y los envían al núcleo SaaS de forma segura, sin exponer datos reales y mediante cifrado TLS.
- **Integración de Identidad:** Conexión con Azure Active Directory (AD) de CEPESA E&P para autenticación unificada (SSO) y sincronización de usuarios y grupos, garantizando una experiencia segura y consistente.

Impacto y Beneficios Clave del Proyecto

Gracias a la solidez de Anjana Data y al acompañamiento experto del Customer Success Manager, junto con la colaboración activa de los equipos implicados, se han alcanzado beneficios concretos que aceleran la transformación digital y fortalecen la gobernanza del dato en CEPSA E&P, como son:

- **Visión 360° del Linaje del Datos:** Conexión de entornos previamente aislados (ficheros locales, procesos en Python, infraestructura Cloud y herramientas BI) en un linaje gráfico navegable y comprensible tanto para perfiles técnicos como de negocio.
- **Glosario de Negocio Avanzado con Navegación Inteligente:** Relaciones semánticas entre términos, KPIs y activos técnicos permiten una experiencia fluida. Un usuario puede pasar de un concepto como “Wellbore” a los informes de Power BI que lo utilizan y a las tablas físicas que lo contienen, obteniendo contexto completo con un solo click.
- **Eficiencia Operativa:** Centralización del conocimiento que elimina la dependencia del “conocimiento tribal”, agiliza la resolución de incidencias y acelera la incorporación de nuevos empleados.
- **Escalabilidad y Evolución del Modelo:** Validación de integraciones con Microsoft Purview DLP y Documentum D2, demostrando que Anjana Data está lista para evolucionar hacia un gobierno integral que abarque datos estructurados, seguridad y gestión documental.