



Por *Juan Coria, Gonzalo Sisto, Rodolfo Figueroa y Marcelo Giaquinta (YPF) y Ricardo Handson (Axonier)*

Un repaso por la estrategia de la operadora argentina para integrar los numerosos archivos físicos y digitales generados a través de su amplia e histórica operatoria.

Planteo del problema

Dentro del ámbito de la información técnica de *upstream*, YPF posee una veintena de archivos físicos, que contienen más de un millón y medio de elementos distribuidos en las distintas operaciones del país, y un Gestor Documental con más de dos millones de documentos digitales.

Por un lado, con respecto a los archivos físicos, no existía una solución integrada y centralizada para gestionarlos conjuntamente, por lo tanto, la catalogación de los elementos carecía de validaciones estándares y dificultaba la implementación de un proceso homogéneo de gestión de préstamos y devoluciones. Esto generaba inconsistencias en la clasificación entre los diferentes archivos técnicos e impedía las consultas, en forma remota, del material en los distintos sitios, lo cual impactaba en la disponibilidad y el uso de la información.

Por otro lado, el Gestor Documental requería una reorganización de los metadatos para soportar nuevas clasificaciones y una actualización de la aplicación con el fin de incorporar nuevas funcionalidades.

En este contexto, se implementó un proyecto para integrar estos ambientes de información y ampliar las

Gestión integrada de activos físicos y digitales

facilidades de búsqueda, explotación y administración del contenido.

Se comenzó con la revisión, reorganización e integración de los metadatos del Gestor Documental, así como también del catálogo de los Archivos Físicos, y se continuó con la búsqueda e implementación de una única plataforma que debía reemplazar a las dos existentes, que estuviera acorde con los avances tecnológicos actuales, que incorporara nuevas funcionalidades y, además, que sea “amigable” para los usuarios finales y los administradores del contenido.

Desarrollo técnico del trabajo

Análisis y redefinición del modelo de metadatos basado en dimensiones

Con el objetivo de clasificar la información estratégica proveniente de los diferentes sectores y nuevas áreas incorporadas a la compañía a través del tiempo, se buscó unificar los criterios de clasificación de la información física y digital, ya que no solo se busca almacenar el contenido de los documentos y los elementos físicos, sino que también se necesita clasificarlos correctamente dentro

de la estructura documental y dotarlos de los metadatos necesarios para garantizar su accesibilidad, integridad y autenticidad.

El proceso requiere definir previamente el modelo de clasificación que, en estos ambientes, se basa en la definición de un conjunto de metadatos, tanto para la identificación de contenido como para el resto de su gestión.

La premisa fue obtener un modelo de clasificación lógica con el propósito de encontrar la información de manera más rápida y eficiente, que no se basara solamente en “navegar” a través de una jerarquía de documentos o datos de un elemento físico, sino que cada elemento o documento debía estar descrito a partir de metadatos atributos integrados, comunes y únicos, que permitan accederlos directamente.

Para el desarrollo del nuevo modelo de metadatos y la migración del actual al mismo, se trabajó en base a los siguientes aspectos:

- ✓ Análisis de las clasificaciones existentes del Gestor Documental, de los archivos físicos y de otros entornos como bases de datos con información relacionada.
- ✓ Usos del lenguaje y nombres locales: los términos históricamente utilizados en YPF fueron tomados en



cuenta –por relevamiento– para ajustar el modelo, ya que a través de ellos se clasifican e interpretan los contenidos y se busca la información por parte de los usuarios.

- ✓ Consolidación del orden, la clasificación y la relación con otros entornos que el modelo existente resolvía: en el modelo del Gestor Documental, cada documento estaba asociado a 65 metadatos, que reflejaban las maneras, los usos y las necesidades de ordenar, clasificar y buscar la información que se fueron utilizando a través del tiempo. Algunos definidos por los administradores y otros necesarios por la lógica de las aplicaciones que los gestionaban.
- ✓ Modelos estándares: se revisaron las definiciones, las metodologías y las sugerencias de estándares de la industria y otros (PPDM, Energistics, Dublin Core, etc.), así como también las de otros proveedores de taxonomías.
- ✓ Requerimientos Legales: establecidos por las autoridades de aplicación (nacionales y provinciales, entes de control, etc.) y procesos del negocio relacionados con dichos requerimientos.
- ✓ Compatibilidad: el nuevo modelo de clasificación debía ser compatible con los modelos de repositorios (BD) existentes, dado que en general los documentos y elementos físicos están asociados con grandes volúmenes de información digital almacenadas en las bases de datos, que tienen sus modelos de metadatos por el que acceden los usuarios.

En función de estos criterios de análisis se generó un nuevo modelo basado en “dimensiones”. El concepto trata de agrupar los metadatos en subconjuntos de la misma “naturaleza”, es decir identificar el sentido con el cual se definió y se utiliza cada metadato. A modo de ejemplo, se identificaron dimensiones para la gestión de los elementos/documentos, para la administración de seguridad, para la asociación con procesos del negocio, para las ubicaciones geográficas y para la relación con otros repositorios, entre otros. Esto simplifica el diseño, el mantenimiento y la extensión del modelo completo. Asimismo, si los elementos de cada subdimensión tienen una relación de jerarquía entre ellos, por cada dimensión permite la creación de “facetas” (*facets*) lo que brinda, si es necesario, distintos “árboles de navegación” en jerarquías especializadas y consistentes, y evita la definición de una única jerarquía para todos los metadatos, que es

muy difícil de lograr sin sacrificar niveles de clasificaciones. En resumen, las facetas permiten que un mismo elemento pueda ser encontrado “navegando” jerarquías de metadatos de distintas naturalezas.

Una de las dimensiones más importante está formada por los metadatos que describen la asociación de los (contenidos de) documentos/elementos con el *upstream*. Para mantener la simplicidad del modelo se dividió la posible información del *upstream* a gestionar en “activos o maestros” por especialidad y de esta manera luego se utiliza metadatos de “tipo y subtipo” específica.

El modelo resultante fue aplicado, soportando la información técnica existente y permitiendo la integración entre los ambientes Documental y de Archivos Técnicos.

Evaluación de arquitecturas y elección de la nueva plataforma técnica

Siguiendo los lineamientos estratégicos de la compañía, se analizaron algunas herramientas disponibles en el mercado que sean capaces de cubrir y resolver los procesos de negocio actuales, cubrir las deficiencias o faltantes y sentar las bases en un punto que permita el crecimiento, la incorporación de nuevos orígenes y una fácil evolución hacia nuevas herramientas y plataformas con la respectiva integración.

Para este fin, se debía integrar, no solo dos procesos claves de *upstream*, sino la información de los archivos físicos que estaba soportada en sistemas *stand-alone*, lo cual implicaba un fuerte trabajo de normalización y posteriormente la migración de información.

Dentro de la evaluación se ponderaron ciertos requisitos de mínima, que fueron los siguientes: una plataforma en la nube, la cual debía ser factible de elegir por YPF, ser líder en la industria, con amplia experiencia a nivel internacional en implementaciones que la envergadura de YPF requiere y que agilice y facilite la integración, tanto en la misma nube como contra sistemas *on-premise* para el intercambio de información, sea API compatible trabajando vía servicios.

El logro de encontrar esta combinación de componentes aseguró estar en un punto óptimo tanto tecnológico como de capacidad de crecimiento para encarar los nuevos desafíos que la industria requiere día a día.

La evaluación consistió en la ejecución de pruebas de concepto de empresas que disponían plataformas con estas características, las cuales fueron evaluadas por espe-





cialistas de las distintas áreas involucradas para analizar su factibilidad, entre las que se destacaron: Tecnología de Operaciones, Data Management, Data Governance, Seguridad, Ciberseguridad, Arquitectura, Infraestructura y Especialidades Técnicas y SSII.

Esta fue una de las primeras implementaciones de aplicaciones de *upstream* de YPF en el mundo de “la nube”, lo cual representó un desafío adicional tecnológicamente y en paralelo, ya que se estaban desarrollando los conocimientos técnicos, la definición de políticas y demás temas implicados en la transformación. Este trabajo brindó un sinnúmero de lecciones aprendidas y experiencias adquiridas para los nuevos desafíos.

Resultados obtenidos

Luego de la ejecución de un proyecto del que participaron una variedad de equipos técnicos, el éxito de este se puede medir a través de la obtención de distintos objetivos planificados, todos ellos alineados para soportar uno en común: el aseguramiento del resguardo y la maximización del uso de la información por parte de los técnicos del *upstream*, a saber:

- Redefinición del modelo de metadatos con la máxima integración entre el Gestor Documental y Archivos Técnicos.
- Alineación del modelo de metadatos con los modelos de los repositorios (BD) con información relacionada.
- Plataforma técnica estable y escalable.
- Flexibilidad y facilidades de integración de la herramienta con resto de sistemas base para el correcto funcionamiento de los flujos de proceso, como así también nuevos requerimientos de valor para otros.
- Nuevas facilidades de los módulos de préstamos y devoluciones para la gestión de la información física de los archivos de YPF.
- Posibilidad de consulta y gestión simultánea a la información digital/ física de múltiples usuarios en diferentes ubicaciones.
- Búsqueda e identificación más precisa y rápida de la

información. OCR de contenido y técnicas de manejo de información estructurada y no estructurada.

- Mejor gestión de intercambio entre las fuentes integradas y los usuarios.
- Simplificación y optimización los procesos de clasificación de la información.
- Modelo de denominaciones más consistentes en la catalogación, lo que permite un acceso más fácil.
- Favorecimiento de la implementación de políticas de destrucción de registros/documentos físicos, reduciendo los costos de almacenamiento físico.
- Permite homogeneizar las políticas de acceso y auditorías de contenido, como también indicadores y trazabilidad de uso.

Conclusiones

Dentro del gran abanico de información y tipos de datos que gestiona YPF en la industria del *upstream*, el cambio de plataforma técnica de la gestión integrada de documentos y elementos físicos en gran volumen y distribuidos geográficamente presentó desafíos no menores.

Resultó fundamental analizar ampliamente la lógica de clasificación que se utilizaría teniendo en cuenta los metadatos no solo como identificación interna de los elementos, sino su relación con el resto de los repositorios y el manteniendo, en lo posible, de la terminología históricamente usada para referirse a la información. Todo esto dentro de un marco que permitiera la extensión del modelo para incorporar nuevos tipos de documentos o elementos que se requieran en el tiempo.

Además, la incorporación de una aplicación basada en la nube heredó los beneficios de acceso, la disponibilidad, las nuevas funcionalidades y el manejo de nuevos tipos de datos que brindan hoy las nuevas tecnologías.

El conjunto de estas mejoras transformó dos ambientes de información técnica que se encontraban separados en uno solo, donde las nuevas facilidades de uso y conectividad entre los usuarios y los administradores lo convierten conceptualmente en un ambiente colaborativo de explotación de la información.