

Control de normativas

A continuación, se expone un cuadro resumen de las normativas actualmente vigentes en el uso de la plataforma Big Data y cómo se está llevando a cabo la implementación de dichas normativas. Además, para aquellas en las que se ha detectado que no existe un mecanismo de control, se realiza una propuesta por parte del equipo de Gobierno del Dato para poder solventarlo.

Norma	Forma de control	Nombre visualización	Propuesta de control	Comentario
Normativa para el paso a producción				
Se cumple con la estructura del activo establecida por la organización	Cuadro de Mando Reglas de Calidad y Cuadro de Mando de Glosario y Ontologías	Estructura Comunidades-Dominios-Términos y Ubicación de las reglas de calidad	-	-
Se ha definido el propietario del dato	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Metadatos no intrínsecos asignados	-	Es un custom attribute que se verá en el Cuadro de Mando correspondiente cuando se defina para algún activo
El nombre cumple con la normativa establecida por la organización	Cuadros de Mando	Apartado de revisión nomenclatura Cuadro de Mando Colecciones, Technical Views y Business Views, Cuadro de Mando de Glosario y Ontologías, Cuadro de Mando Reglas de calidad y Cuadro de Mando Rocket	-	-
El campo de descripción del activo se encuentra completo	Cuadros de Mando	Descripción Colecciones, Descripción Technical Views, Descripción de los ítems (Cuadro de Mando de Glosario y Ontologías), Descripción de Reglas de Calidad, y Descripción de los proyectos y Assets sin descripción (Cuadro de Mando Rocket)	-	-
Los permisos de acceso y consumo han sido definidos y validados	GoSec	-	-	La herramienta incluye un apartado de auditorías donde se ve quién y a qué ha accedido/modificado cada usuario
Contiene los metadatos mínimos asociados	Cuadros de Mando	Apartados específicos en Cuadro de Mando Colecciones, Technical Views y Business Views, Cuadro de Mando de Glosario y Ontologías, Cuadro de Mando Reglas de Calidad y Cuadro de Mando Rocket	-	-
El formato de los metadatos asociados es el correcto	Cuadros de Mando	Apartados específicos en Cuadro de Mando Colecciones, Technical Views y Business Views, Cuadro de Mando de Glosario y Ontologías, Cuadro de Mando Reglas de Calidad y Cuadro de Mando Rocket	-	-
La sensibilidad de la información ha sido definida y validada	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Apartado "Atributo sensibilidad" Cuadro de Mando Colecciones, Technical Views y Business Views	-	Es un metadato de las columnas de las Technical Views

Las reglas de calidad han sido aprobadas por el responsable correspondiente	Cuadro de Mando Reglas de Calidad	Responsable de la Regla de Calidad	-	En el Cuadro de Mando indicado puede verse el estado de las reglas, pero no quién la ha aprobado
Se ha definido y validado el umbral necesario para las diferentes reglas de calidad	Cuadro de Mando Reglas de Calidad	Umbrales	-	-
Los resultados de los perfilados de datos son tratados correctamente	Cuadro de Mando Reglas de Calidad	Tratamiento de los resultados de los perfilados de datos	-	-
La frecuencia de ejecución de las reglas de calidad ha sido definida y validada por el responsable correspondiente	Cuadro de Mando Reglas de Calidad	Frecuencia de ejecución válidas	-	-
La tipología de relaciones utilizada está dentro de las establecidas por la organización	Cuadro de Mando Relaciones	Tipos de relaciones entre assets	-	-
La tipología de activos de negocio utilizados se recoge en las acordadas por la organización	Cuadros de Mando	-	-	Es algo intrínseco controlado en los Cuadros de Mando
Los términos de negocio han sido definidos y validados por el responsable acordado.	Cuadro de Mando de Glosario y Ontologías	Custom Attributes por Business Terms	-	En el Cuadro de Mando de Glosario y Ontologías se puede ver el responsable, pero no si ha sido aprobado por él o por otro usuario
El linaje de los datos ha sido identificado y definido por el responsable correspondiente	Cuadro de Mando Linaje	Origen y transformación	-	-
Normativa para el control de versiones				
Normativa para el control de versiones	No	-	Se propone una supervisión manual	-
Normativa para definición de linaje				
Normativa para definición de linaje	Cuadro de Mando Linaje	Cuadro de Mando Linaje	-	-
Circuito de subida a producción				
Circuito de subida a producción	No	-	Se propone una supervisión manual	-
Requisitos mínimos de metadatos				

Metadatos comunes de Governance	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Metadatos intrínsecos nulos y Metadatos intrínsecos nulos – Technical Views	-	-
Metadatos específicos de colecciones	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Apartado de revisión de metadatos mínimos de Colecciones	-	-
Metadatos específicos de Technical Views	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Apartado de revisión de metadatos mínimos de Technical Views	-	-
Metadatos específicos de Business Views	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Apartado de revisión de metadatos mínimos de Business Views	-	-
Metadatos específicos de Reglas de Calidad	Cuadro de Mando Reglas de Calidad	Apartado de revisión de metadatos mínimos del Cuadro de Mando de Reglas de Calidad	-	-
Metadatos específicos de Ontologías	Cuadro de Mando de Glosario y Ontologías	Custom attributes asignados a elementos de ontologías	-	-
Metadatos específicos de Comunidades	Cuadro de Mando de Glosario y Ontologías	Apartado de revisión de metadatos mínimos de Comunidades	-	-
Metadatos específicos de Dominios	Cuadro de Mando de Glosario y Ontologías	Apartado de revisión de metadatos mínimos de Dominios	-	-
Metadatos comunes de Rocket	Cuadro de Mando Rocket	Apartado de revisión de metadatos mínimos del Cuadro de Mando de Rocket	-	-
Nombramiento de Assets				
Nombramiento de Reglas de Calidad	Cuadro de Mando Reglas de Calidad	Apartado de revisión de nomenclatura del Cuadro de Mando de Reglas de Calidad	-	-
Nombramiento de Technical Views	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Revisión de nomenclatura de Technical Views	-	-
Nombramiento de Colecciones	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Revisión de nomenclatura de colecciones	-	-

Nombramiento de Constantes	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Listado de constantes	-	No hay constantes definidas actualmente por usuarios, solo las que trae por defecto la herramienta.
Nombramiento de nuevos atributos adicionales	Cuadro de Mando Colecciones, Technical Views, Business Views y nombramiento de custom attributes	Nombramiento de Custom Attributes	-	Los nuevos custom attributes creados deben nombrarse como atr_[Atributo adicional] y esto no se está controlando
Nombramiento de Comunidades, Dominios y Términos de Negocio	Cuadro de Mando de Glosario y Ontologías	Comunidades definidas y Dominios definidos	-	Habrà que modificarlo porque serán los que se definan en el Mapa Informacional de Sistemas
Nombramiento de Ontologías	Cuadro de Mando de Glosario y Ontologías	Nombres de ontologías que no cumplen la normativa	-	-
Nombramiento de proyectos de Rocket	Cuadro de Mando Rocket	Apartado de revisión de nomenclatura de los proyectos	-	-
Nombramiento de directorios de Rocket	Cuadro de Mando Rocket	Nombres de directorios que no cumplen la normativa de nomenclatura	-	-
Nombramiento de assets de Rocket	Cuadro de Mando Rocket	Assets sin Orden de Ejecución ni control de versionado y Assets que no cumplen la normativa de nomenclatura	-	-
Normativa de documentación				
Nuevas fuentes de datos	No	-	Se propone una supervisión manual	-
Reglas de Calidad	No	-	Se propone una supervisión manual	-
Normativa de gestión de assets				
Gestión de relaciones – Relaciones entre activos	Cuadro de Mando Relaciones	Tipos de relaciones entre assets	-	-
Gestión de relaciones – Tipos de relaciones	Cuadro de Mando Relaciones	Tipos de relaciones entre assets	-	-
Normativa de Airflow				
Nomenclatura de DAGs	Normativa y PID	-	-	-
Etiquetas DAGs	Normativa y PID	-	-	-
Estructura de carpetas de DAGs	Normativa y PID	-	-	-

Estructuras de carpetas de Librerías	Normativa y PID	-	-	-
Directrices de llamadas a flujos	Normativa y PID	-	-	-
Documentación de DAGs	Normativa y PID	-	-	-
Parámetros de DAGs y grupos de tareas	Normativa y PID	-	-	-
Normativa de ficheros parquets				
Los parquets particionados no deben pesar más de 1 GB ni menos de 250 MB	Cuadro de mando de Parquets	Tamaño parquets	-	-
La columna de partición no contiene nulos y es equidistante	Cuadro de mando de Parquets	La columna de partición es equidistante	-	-
Se han ingestado y curado aquellas tablas que requerían añadir alguna columna	Normativa y PID	-	-	-
Si la tabla tiene menos de 5 millones de filas, se almacena en Postgres	Normativa y PID	-	-	-
Las tablas ingestadas que van a HDFS se incluyen en una carpeta llamada "dev/raw" o "dev/raw_incremental" del entorno DEV, o "raw" o "raw_incremental" de los entornos PRE o PRO	Cuadro de mando de Parquets	Capas	-	-
Las tablas curadas en HDFS se incluyen en una carpeta llamada "curado_sas" o "curado_sas_incremental"	Cuadro de mando de Parquets	Capas	-	-
Los datos ya ingestados y curados pasan a la capa de consumo, cuya carpeta en HDFS es "consumo" o "consumo_incremental"	Cuadro de mando de Parquets	Capas	-	-
Se realizan periódicamente copias de seguridad de las carpetas "..._incremental"	Normativa y PID	-	-	Se guarda en la carpeta backup de HDFS
Normativa Postgres				
Las particiones, índices y claves primarias deben ser idénticos en Postgres y en Oracle	Normativa y PID	-	-	-
Las tablas almacenadas en Postgres tienen menos de 5 millones de registros	Cuadro de mando de Postgres	Tablas con más de 5 millones de registros	-	-
Las tablas son consistentes en todas las instancias	Normativa y PID	-	-	-
Los nombres de las tablas, campos, secuencias, índices, constantes y esquemas siguen las pautas establecidas en la normativa	Cuadro de mando de Postgres	Tablas que no tiene prefijo ni sufijo correcto, Tablas cuyo nombre contiene algún carácter en minúscula, Nombres de columnas que no tienen prefijo correcto, Nombres de columnas que contienen algún carácter en minúscula, Nombres de índices que tienen algún carácter en minúscula, Nombres de constantes que tienen algún carácter en minúscula	-	-
Normativa parametrización Rocket				

Se han establecido los parámetros necesarios en cada proyecto	No	-	Se propone una supervisión manual	-
Se han activado y usado los grupos de parámetros necesarios en cada workflow	No	-	Se propone una supervisión manual	-
Normativa de desarrollos de Rocket				
Se siguen las buenas prácticas impuestas para los desarrollos de Rocket	Normativa y PID	-	-	-