



# Informe Técnico: Revisión de Dominio de Datos – RESULTADOS



## Contenido

<b>1. Objetivos.....</b>	<b>3</b>
Objetivo general .....	3
Objetivos específicos.....	3
<b>2. Resultados.....</b>	<b>4</b>
2.1. Actualización de vistas materializadas .....	4
2.2. Consulta de dominios definidos VS Consulta de dominios usados en los datos.....	4
2.3. Comparación de dominios (reporte de inconsistencias) .....	5
<b>CONCLUSIÓN.....</b>	<b>7</b>

## Índice de tablas

Tabla 1: Resultado de dominios parametrizados en arcgis .....	5
Tabla 2: Resultado de dominios iniciales insertados .....	5
Tabla 3: Resultados individuales que no corresponden a los dominios en ArcGIS .....	6
Tabla 4: Tabla de resultados generales de inconsistencias entre dominio esperado e insertado .....	7

## Índice de Ilustraciones

Ilustración 1: Actualización de vistas para la estructura de dominios esperados y los insertados .....	4
Ilustración 2: Resultado de la comparación entre valores usados y dominios definidos (inconsistencias detectadas) .....	6

## 1. Objetivos

### Objetivo general

Garantizar la **integridad y consistencia** de los valores de dominio configurados en ArcGIS con respecto a los datos efectivamente almacenados en la base de datos colsmart\_preprod\_migra, mediante la actualización de vistas materializadas y la detección automática de discrepancias.

### Objetivos específicos

- ✓ **Reconstruir la estructura completa** de los dominios definidos en la GDB, incluyendo los valores válidos asociados a cada columna.
- ✓ **Almacenar de forma persistente** tanto la estructura esperada como los datos utilizados mediante vistas materializadas para facilitar consultas posteriores.
- ✓ **Permitir comparaciones automáticas** entre los valores usados y los definidos en los dominios para detectar inconsistencias, errores de digitación u homologaciones pendientes.
- ✓ **Facilitar la trazabilidad** entre la implementación de modelos de datos espaciales y el uso efectivo de las reglas de negocio asociadas a los dominios en el sistema de gestión de base de datos.
- ✓ **Generar reportes detallados de inconsistencias** que prioricen los casos críticos y sirvan de insumo para planes de depuración y homologación de datos.
- ✓ **Integrar el proceso de validación de dominios** dentro de los flujos ETL y las tuberías CI/CD, garantizando verificaciones recurrentes antes de cada despliegue o migración.

## 2. Resultados

### 2.1. Actualización de vistas materializadas

El primer paso del script consiste en la actualización manual de dos vistas materializadas: `colsmart_preprod_migra.estructura` y `colsmart_preprod_migra.estructura_data`. Esta acción se realiza mediante el comando SQL `REFRESH MATERIALIZED VIEW`, el cual fuerza la recarga de datos desde las fuentes originales para que ambas vistas reflejen el estado más reciente de la base de datos.

La vista `estructura` contiene la definición teórica de los dominios configurados en ArcGIS, es decir, la relación entre tablas, columnas y los valores válidos permitidos por cada dominio. Esta información se extrae directamente desde el XML almacenado en la Geodatabase (GDB), y representa la configuración de referencia que deberían cumplir los datos reales (ver **DOCUMENTO\_CALIDAD.pdf**, págs. 7–8). Por su parte, la vista `estructura_data` almacena los valores efectivamente insertados en las tablas de la base de datos para esos mismos campos con dominio, generados dinámicamente mediante la función `distinct_por_columna()` (ver **DOCUMENTO\_CALIDAD.pdf**, págs. 5–6). Estas vistas materializadas permiten realizar comparaciones automáticas entre lo definido como válido y lo que realmente se usa, y fueron desarrolladas expresamente para apoyar procesos de validación y control de calidad de datos espaciales en entornos ArcGIS y PostgreSQL.

Dado que las vistas materializadas no se actualizan automáticamente, es fundamental refrescarlas antes de realizar cualquier verificación o análisis de inconsistencias. Si se han cargado nuevos datos, modificado dominios o ejecutado procesos ETL, esta actualización asegura que el análisis posterior se base en información confiable y actualizada. No ejecutar este paso puede llevar a diagnósticos erróneos, como identificar errores que ya fueron corregidos o pasar por alto problemas recientes. Por lo tanto, refrescar estas vistas constituye una etapa crítica del procedimiento de validación de dominios.

```

****
* Se debe refrescar la vista para actualizar la estructura de dominios
*/

REFRESH MATERIALIZED view colsmart_preprod_migra.estructura;

****
* Se debe refrescar la vista para actualizar los datos insertados en la estructura de dominios
*/

REFRESH MATERIALIZED view colsmart_preprod_migra.estructura_data;

```

Ilustración 1: Actualización de vistas para la estructura de dominios esperados y los insertados

### 2.2. Consulta de dominios definidos VS Consulta de dominios usados en los datos

Una parte esencial del proceso de validación consiste en contrastar los **dominios definidos en la Geodatabase de ArcGIS** con los **valores realmente utilizados en la base de datos**. Esta comparación permite detectar errores comunes como valores no válidos, errores de digitación, o dominios mal configurados. Para ello, el script implementa dos consultas distintas sobre vistas

materializadas: una que representa la estructura oficial de los dominios (estructura) y otra que refleja los datos operativos (estructura\_data). A continuación, se presenta una comparación entre ambas consultas para ilustrar claramente su propósito, fuente y utilidad dentro del proceso de aseguramiento de calidad.

Tabla 1: Resultado de dominios parametrizados en arcgis

RESULTADO ESPERADO			
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Bueno
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Deficiente
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Intermedio
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Malo
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_Bueno
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_Malo
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Optimo
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Regular
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Ruinoso
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Optimo_1

VS

Tabla 2: Resultado de dominios iniciales insertados

RESULTADO OBTENIDO			
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Malo_4
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Regular_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_bueno_1_5
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_Bueno_1_5
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Bueno_2
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_bueno_1_5

### 2.3. Comparación de dominios (reporte de inconsistencias)

La tercera etapa del proceso consiste en comparar los valores utilizados en la base de datos con los valores válidos definidos en los dominios de ArcGIS. Esta comparación permite generar un **reporte de inconsistencias**, es decir, un listado de valores que **han sido insertados en los datos pero no están definidos en el dominio correspondiente**.

Este bloque compara campo por campo y tabla por tabla, devolviendo únicamente aquellos valores que aparecen en los datos pero que no están contenidos en el dominio oficial. También permite aplicar filtros opcionales (comentados) para acotar el análisis a una tabla o columna específica. El resultado de esta consulta es una tabla de inconsistencias que se puede usar como **insumo para tareas de limpieza y homologación**, o como evidencia en auditorías de calidad de datos. Esta técnica de validación cruzada es altamente eficiente, especialmente cuando se trabaja con volúmenes grandes de datos, ya que evita el uso de subconsultas complejas y permite una ejecución directa sobre vistas materializadas optimizadas.

Tal como se describe en el *DOCUMENTO\_CALIDAD.pdf* (págs. 8–9), esta comparación se apoya en la persistencia previa de las vistas estructura y estructura\_data, lo cual permite detectar errores de forma ágil, reproducible y fácilmente integrable en procesos ETL o validaciones automáticas.

A? schema_name	A? table_name	A? column_name	A? name_value
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Bueno_2
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_bueno_1_5
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Malo_4
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Regular_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_bueno_1_5
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	conservacion_tipologia	Muy_Bueno_1_5
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	5021143_Institucional.Tipo_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	2021132_Comercial.Intermedio_1
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	4024022_Conservacion.Residencial_Sencilla_Tipo_2
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	6021131_Institucional.Religioso_Tipo_1
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	2021532_Comercial.Intermedio_2
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	Anexos.Coliseos_Plus_Tipo_80
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	9016194_ED.Multifamiliar_VIS_Hasta_12_Pisos
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	Anexos.Plazas_de_Toros_Concreto_Tipo_80
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	1021134_Residencial.Tipo_4
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	5021132_Institucional.Tipo_7
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	1024114_Residencial.Tipo_4_Menos
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	2026532_Comercial.Intermedio_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	5036553_Institucional.Tipo_8
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	3011132_Industrial.Mediana_1
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	4031035_Conservacion.Residencial_Tipo_5_Restaura
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	5011111_Institucional.Tipo_6
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	7031173_Institucional.Salud_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	2033133_Comercial.Especializado_3
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	1011133_Residencial.Tipo_3_mas
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	9026547_ED.Servicios_Tipo_1
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	4014011_Conservacion.Residencial_Sencilla_Tipo_1
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	5011122_Institucional.Tipo_2
colsmart_preprod_migra	ilc_caracteristicasunidadconstruccion	tipo_tipologia	2023123_Comercial.Especializado_1

**Ilustración 2:Resultado de la comparación entre valores usados y dominios definidos (inconsistencias detectadas)**

**Tabla 3: Resultados individuales que no corresponden a los dominios en ArcGIS**

tipo_tipologia - ilc_caracteristicasunidadconstruccion
5021132_Institucional.Tipo_7
1021125_Residencial.Tipo_5
5011122_Institucional.Tipo_2
2014111_Comercial.Basico_2
.....

tipo - ilc_fuenteadministrativa
SENTENCIA
EXTRACTO
ACTA DE CONCILIACION
DESPACHO COMISORIO
.....

tipo-ilc_predio
Departamental
Municipal
Nacional
Baldio
.....

conservacion_tipologia - ilc_caracteristicasunidadconstruccion
Malo_4
Regular_3
Muy_bueno_1_5
Muy_Bueno_1_5
.....

uso - ilc_caracteristicasunidadconstruccion
comercial_cimientos, estructura, muros y placa base centros comerciales en ph (lorenzo-pasto)
Residencial_No_Especificado

grupo_etnico - ilc_interesado
Etnico_Indigena
Etnico_Negro_Afrocolombiano

tipo_documento -ilc_interesado
Libreta Militar
Otro

tipo -ilc_derecho
Sin definir

**Tabla 4: Tabla de resultados generales de inconsistencias entre dominio esperado e insertado**

Tabla Resumen de errores		
ilc_interesado	porcentaje_propiedad	93
ilc_caracteristicasunidadconstruccion	tipo_tipologia	53
ilc_fuenteadministrativa	tipo	24
ilc_predio	tipo	11
ilc_caracteristicasunidadconstruccion	conservacion_tipologia	6
ilc_caracteristicasunidadconstruccion	uso	2
ilc_interesado	grupo_etnico	2
ilc_interesado	tipo_documento	2
ilc_fuenteadministrativa	ia	1
ilc_interesado	autorreco_campesino	1
ilc_interesado	nombre_pueblo	1
ilc_derecho	tipo	1
ilc_predio	condicion_predio	1
Total		198

## CONCLUSIÓN

Los resultados obtenidos a través del script de validación de dominios permiten evidenciar de forma clara y estructurada la coherencia —o falta de ella— entre los valores definidos como válidos en la Geodatabase de ArcGIS y los datos efectivamente registrados en la base de datos. La consulta sobre la vista estructura proporciona una referencia normativa confiable, mientras que estructura\_data ofrece una fotografía actual de los valores en uso. Al comparar ambas mediante un EXCEPT, se logra identificar de forma precisa aquellos registros que no cumplen con las reglas del modelo conceptual, revelando inconsistencias, errores de digitación, o ausencias en la configuración de dominios.

Este enfoque no solo permite detectar errores existentes, sino también anticipar y prevenir conflictos futuros en procesos ETL, validaciones automáticas o migraciones. La persistencia de los datos en vistas materializadas optimiza la consulta y facilita la integración de estos controles en flujos automatizados. En consecuencia, el sistema implementado se constituye como una herramienta clave para garantizar la calidad y trazabilidad de los datos espaciales, reforzando la alineación entre el diseño del modelo y su implementación operativa.