# FASE N° 2.

## SISTEMA UNIFICADO DE INFORMACIÓN DE LOS 3 SERVICIOS CONSULTADOS

# Producto:

## 2.1 Recopilación de la información primaria y secundaria de los usuarios de servicios públicos domiciliarios de acueducto, alcantarillado y aseo en el municipio de Florencia

**SISTEMA DE INFORMACIÓN UNIFICADO PARA SEGUIMIENTO Y CONTROL DE LOS SERVICIOS: FASE 2.1**

| Tabla de contenido de Anexos informativos suministrados | |
|---|---|
| **Ubicación Carpeta** | **Nombre de Archivos** |
| **Catastro** | Cancelados_Sector_04.<br>CanceladosyCambiosNPN_Sector_01_02_03.<br>REGISTRO_1. |
| **Control Urbano** | RESOLUCIONES 2023.<br><br>RESOLUCIONES. |
| **Estratificación** | 1. BASE DATOS DECRETO 0190 AÑO 2000 base del 2000.<br><br>BASE DE DATOS ESTRATIFICACION DECRETO 00172 del 27 ABRIL DE 2023..<br><br>LISTADO DE LADOS DE MANZANAS ESTRATIFICADOS. |
| **Secretaria Hacienda** | **Servicios Públicos Domiciliarios:**<br>Cartilla municipal de los servicios públicos domiciliarios<br>concepto_superservicios_0000043_2022<br>concepto_superservicios_0000057_2022<br>CptoServPub<br>Ley_142_de_1994<br>Superservicios-publicos-domiciliarios-en-Colombia-memoria-retos-y-vision<br>- **análisis Cobro de Subsidios:**<br>**- Base Datos 2018:**<br>1. SubsidiosContribuciones_Vigencia_2025_Mes_1_SERVAF ENERO 2025.<br><br>BASE DE DATOS ESTRATIFICACION DECRETO 00172 del 27 ABRIL DE 2023.<br>Predios Excluidos Final.<br>**- ESAC:**<br>2. ESAC-S-202508026 _ Facturación 2024 para liquidación concurso económico 2025.<br>Acuerdo Mcpal N° 2020022 - 2020 Modificacion % Subsidios y Contribuciones 5 años AAA. |

| | |
|---|---|
| | ANALISIS ENERO ESAC.<br>análisis estratos subsidios spd.<br>Balance S&C ESAC enero 2025 V2.<br>BASE DE DATOS ESTRATIFICACION DECRETO 00172 del 27 ABRIL DE 2023.<br>ESAC- S-202508035 _ S&C enero 2025 _ Corrección.<br><br>**- SERVAF NOVIEMBRE 2024:**<br>1. BASE DATOS A DICIEMBRE 2022 – VIGENTE.<br><br>ANALISIS SERVAF DE NOVIEMBRE 2024.<br><br>BasedeDatosELECTROCAQUETA-COMERCIAL-INDUSTRIAL-OFICIAL-Consumo Dic Facturacion Ene 2025.<br><br>BasedeDatosSERVINTEGRALDICIEMBRE 2024.<br><br>CRUCE BASE DE DATOS ESTRATIFICACION DECRETO 00172 - SERVICIOS PUBLICOS. |
| **Servicio Publico** | **Aguas de Florencia:** AGUAS DE FLORENCIA JUNIO 6. BASE DE DATOS SUBS ALCANTARILLADO MES JUNIO-25.<br><br>**ESAC:** Balance S&C ESAC junio 2025.<br><br>**SERVAF:** SERVAF JUNIO-6. SubsidiosContribuciones_Vigencia_2025_Mes_6_Ciclo_1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16.<br><br>**SERVINTEGRAL:** SERVINTEGRAL-1830-ANEXO-TOTAL JUNIO 2025. |
| **Super Intendencia** | 20254002865331 (1).<br><br>Anexo 1 20254202780171.<br><br>Anexo 2 20254252805401.<br><br>Estudio tarifario ESAC-2011- Aprob-SSPD.<br><br>Estudio tarifario SERVINTEGRAL 2011-Aprob-SSPD. |

| | |
|---|---|
| | Gmail - Respuesta radicado 20255293483952 del 25 de agosto de 2025_ REPORTE SSPD- facturación_SAMAMB-2024.<br>**ANEXOS RPTA SSPD 05 09 25:**<br>24. Contratos de concesión y operación -20250906T141518Z-1-001 Anexos_Florencia_Planeacion-20250906T141708Z-1-001.<br><br>Punto 28. SERVINTEGRAL PPSA.<br><br>Punto 38.Toneladas de ingreso al sitio disposición relleno sanitario.<br><br>Punto 42. Estados financieros (situación financiera y resultados) 2023 y 2024, -20250906T141520Z-1-001.<br><br>Puntos 23,25,26,36 y 37-20250906T141522Z-1-001.<br><br>**Solicitud habilitación indicadores en SUI:**<br>20254002865331 _1_ _1_ (1).<br><br>Exposición de motivos y acuerdo municipal 160925.<br><br>Gmail - Fwd_ Notificación electrónica radicado entrada No 20255293713242.<br><br>Solicitud de habilitación de indicadores en el aplicativo SUI_INSPECTOR (2). |
| **Super Servicios** | **Acuerdos:**<br><br>**SUI Alcaldía:**<br>EmpresasFla202503312025286081.<br>evaluacion_integral_florencia_vf.<br>evaluacion_integral_servicios_integrales_efectivos_s.a._e.s.p.<br>Reporte SUI 2002 a 2025. |

1. **Recepción y almacenamiento seguro de archivos CSV proporcionados por cada entidad prestadora y verificación del formato, estructura y calidad de los archivos**

**Acciones y acuerdos**

| Acciones | Acuerdos |
|---|---|

| | |
|---|---|
| Entrega de bases de datos por parte de **Licenciamiento (Excel 2019–2025)**. ☑ Cumplido | RESOLUCIONES 2023.xlsx<br>RESOLUCIONES.xlsx |
| Entrega de archivo **R1 de Catastro** con estratificación y capas diferenciadas. ☑ Cumplido | REGISTRO_1.xlsx |
| Almacenamiento seguro de archivos CSV proporcionados por cada entidad prestadora ☑ Cumplido | AGUAS DE FLORENCIA JUNIO 6. BASE DE DATOS SUBS ALCANTARILLADO MES JUNIO-25.xlsx<br>Balance S&C ESAC Junio 2025.xlsx<br>SERVAF JUNIO-6.<br>SubsidiosContribuciones_Vigencia_2025_Mes_6_Ciclo_1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16.xlsx<br>SERVINTEGRAL-CE-GC-1830-ANEXO 4.1-TOTAL JUNIO 2025.xlsx |
| Definición de estructura y formato de archivos en mesa técnica con Catastro. ☑ Cumplido | CanceladosyCambiosNPN_Sector_01_02_03.xlsx<br>Cancelados_Sector_04.xlsx |
| Solicitud de información a operadores de servicios públicos (pendiente armonización completa). | ⚠ Pendiente |

2. **Limpieza y normalización de los datos (campos, nombres, códigos)**

**Acciones y acuerdos**

- Se identificaron inconsistencias: registros municipales (35.000 predios) vs. reportes de operadoras (50.000 usuarios).
- Compromiso de **avanzar en estandarización de datos entre dependencias y operadores**. ☑ Cumplido
- Adopción de la **Ficha Catastral como identificador único**.

- Necesidad de homologar claves catastrales (20 a 30 dígitos). ⚠ En proceso (no completamente cerrado).

## 3. Desarrollo de las Reuniones Técnicas

### 3.1 Mesa Técnica Alcance SIUSP – Secretaría de Planeación

**Fecha y lugar:** 19 de agosto de 2025, Secretaría de Planeación Municipal – Florencia.
**Participantes principales:** Secretaría de Planeación, Oficina de Estratificación, Control Urbano, CORPOAMOR, Catastro Multipropósito.

**Contextualización:** La reunión inició con la presentación de los asistentes y la contextualización por parte del equipo de la consultoría CORPOAMOR. El director José Erley Sandoval explicó que el Sistema Unificado de Información (SIUSP) es uno de los productos centrales de la consultoría y su propósito es fortalecer las capacidades de la Unidad de Servicios Públicos para la toma de decisiones basadas en información integrada y validada.

El coordinador de la consultoría, Iván Salazar, resaltó que el sistema se diseñará con capacidad de interoperar con Catastro Multipropósito, Control Urbano y Estratificación, permitiendo a todas las dependencias acceder a información unificada, actualizada y validada. Durante la socialización del mapa mental del sistema, se explicó cómo este se articulará con planes municipales como el PGIRS, PSMV y PUEAA, y con el Plan de Ordenamiento Territorial vigente.

El equipo de Planeación compartió cifras claves, como los **63.800 predios urbanos y 45.000 subsidios de servicios públicos**, señalando que los datos de operadores no siempre coinciden con la realidad municipal. Surgieron preguntas jurídicas sobre la posibilidad de identificar servicios por ficha catastral, visibilizar licencias nuevas y garantizar la legalidad en la asignación de estratos.

Se enfatizó que la validación de datos será responsabilidad tanto de las empresas prestadoras como de la administración, y que el sistema permitirá identificar desfases y evitar subsidios mal aplicados a usuarios no estratificados. Se concluyó que antes de vincular a los operadores era necesario consolidar la información institucional interna.

**Temas tratados:**

- Presentación del alcance del **SIUSP** como producto de la consultoría contratada.
- Integración del Sistema de información unificado para seguimiento y control de los servicios con catastro, control urbano y estratificación.
- Socialización de un mapa mental del sistema, incluyendo beneficios y funcionalidades.
- Discusión sobre interoperabilidad, acceso a la información por ficha catastral y validación de datos.

**Acuerdos y compromisos:**

- Programar mesas de trabajo con cada dependencia de Planeación (Estratificación, Control Urbano, Catastro Multipropósito). (Cumplido)
- Suministrar la información institucional solicitada para la construcción de la base de datos del sistema. (Cumplido)
- La consultoría entregará manual de usuario y equipo de soporte para el manejo de la herramienta. (Pendiente)

**Observaciones:**
Se destacó la importancia de contar con información validada y actualizada, así como de articular esfuerzos interinstitucionales para evitar desfases entre los datos oficiales del municipio y los reportados por los operadores de servicios públicos.

**Evidencias:**




**3.2 Mesa Técnica Catastro Multipropósito – Sistema de información unificado para seguimiento y control de los servicios**

**Fecha y lugar:** 22 de agosto de 2025, Secretaría de Planeación Municipal – Florencia.

**Participantes principales:** Catastro Multipropósito, Unidad de Servicios Públicos de Planeación, CORPOAMOR.

**Contextualización:** El coordinador Iván Salazar abrió la sesión resaltando la necesidad de unificar la información de la Alcaldía y de Catastro Multipropósito como insumo esencial para nutrir el SIUSP.

Se hizo una **revisión de casos puntuales de estratificación**, lo cual evidenció la importancia de contar con datos homogéneos. El equipo de Catastro planteó inquietudes sobre la conexión entre el sistema y las bases de datos catastrales, el tipo de archivos que debían entregarse y los protocolos de acceso. CORPOAMOR explicó que actualmente se entregan archivos mensuales, pero que era necesario definir una estructura y formato estandarizados para garantizar la interoperabilidad.

La reunión derivó en acuerdos operativos concretos: se programó una mesa técnica el 27 de agosto para definir formatos de archivos y otra con la Unidad de Servicios Públicos para revisar procesos de información de empresas operadoras.

**Temas tratados:**

- Revisión de casos puntuales de estratificación y validación de información con catastro.
- Necesidad de unificar datos de diferentes fuentes para alimentar el Sistema de información unificado para seguimiento y control de los servicios.
- Procedimientos de entrega de archivos e información a la consultoría.

**Acuerdos y compromisos:**

- Realizar una reunión técnica en Catastro el **27 de agosto de 2025** a las 09:00 a.m. para definir estructura y formato de archivos. (Cumplido)
- Agendar reunión técnica con la Unidad de Servicios Públicos el **25 de agosto de 2025** a las 08:00 a.m. para revisar procesos de información de empresas de servicios públicos. (Cumplido)

**Observaciones:**

Se planteó la importancia de definir protocolos claros de intercambio de información, con el fin de garantizar la interoperabilidad entre el **SIUSP** y los sistemas de catastro.

**Evidencias:**





### 3.3 Mesa Técnica SIUSP – Estratificación

**Fecha y lugar:** 25 de agosto de 2025, Secretaría de Planeación – Florencia.
**Participantes principales:** Equipo de Estratificación, Planeación Municipal, CORPOAMOR.

**Contextualización:** El encuentro abordó a fondo el esquema tarifario de acueducto y la aplicación de subsidios y aportes solidarios según la Ley 142 de 1994. Se destacó que los estratos 1 y 2 reciben

subsidios en ambos cargos, el estrato 3 solo en el cargo fijo, y los estratos 4, 5 y 6 pagan tarifa plena, mientras que usuarios comerciales e industriales financian los subsidios.

Se identificó como uno de los mayores problemas la falta de estandarización en la información. Cada empresa operadora maneja bases de datos en diferentes formatos (Excel, TXT, FoxPro), lo que genera inconsistencias y dificulta la trazabilidad. A esto se suma la brecha entre registros municipales (35.000 predios) y los reportes de operadoras (más de 50.000 usuarios).

El grupo analizó la existencia de dos bases de datos principales: una antigua, estructurada por manzanas, y otra más reciente en Excel, con inversión financiera previa, que debe integrarse y actualizarse para mantener la trazabilidad histórica.Se advirtió además que algunas empresas están estratificando por su cuenta en zonas no formalizadas, lo que podría traer problemas jurídicos al municipio. Como solución, se propuso la creación de mesas técnicas permanentes de servicios públicos, con participación de operadores, administración, DANE y Superintendencia.

**Temas tratados:**

- Esquema tarifario de acueducto (cargos fijos y variables) y aplicación de subsidios según Ley 142 de 1994.
- Identificación de aportes solidarios de usuarios comerciales e industriales.
- Falta de estandarización en la información de las empresas operadoras (diferentes formatos de bases de datos).
- Diferencias entre registros municipales (35.000 predios) y reportes de operadoras (50.000 usuarios).
- Necesidad de adoptar la ficha catastral como identificador único.

**Acuerdos y compromisos:**

- Avanzar en la estandarización de datos entre dependencias y operadores. (Cumplido)
- Adoptar la ficha catastral como identificador único de información. (Cumplido)
- Definir formatos claros de información para unificación de bases de datos. (Cumplido)
- Garantizar trazabilidad histórica de la información. (Cumplido)

**Observaciones:**
Se advirtió que la falta de articulación institucional genera desorden en la prestación de servicios y riesgos jurídicos. Se recomendó formalizar mesas técnicas permanentes de servicios públicos con apoyo del DANE y la Superintendencia.

**Evidencias:**




### 3.4 Reunión Técnica SIUSP – Licenciamiento – Control Urbano

**Fecha y lugar:** 26 de agosto de 2025, Oficina de Licenciamiento – Secretaría de Planeación Municipal.

**Participantes:** Asesores de Planeación, CORPOAMOR, contratistas.

**Contextualización:** El ingeniero Iván Salazar explicó la necesidad de articular Catastro, Estratificación y Licenciamiento en el marco del SIUSP para enfrentar la problemática de prestación irregular de servicios en zonas no formalizadas. Señaló que las empresas operadoras han contribuido al desorden al prestar servicios en predios sin legalización.

El asesor Wilson Orlando Bello detalló el procedimiento de licenciamiento: verificación legal de predios, asignación de matrícula inmobiliaria y generación de fichas catastrales. Explicó cómo, a partir de un predio madre, se individualizan los lotes, se asignan áreas remanentes para uso público y se formaliza jurídicamente la construcción.

Se destacó que la información de Catastro es vital para garantizar la legalización predial y que sin licenciamiento adecuado se corre el riesgo de mantener subsidios en áreas ilegales.

**Temas tratados:**

- Necesidad de integrar Catastro, Estratificación y Licenciamiento en el marco del SIUSP.
- Procedimientos de verificación legal aplicados en la formalización de predios.
- Problemática derivada de la prestación irregular de servicios públicos en zonas no formalizadas.

**Compromisos:**

- Enviar base de datos en Excel 2019–2025 (Licenciamiento). (Cumplido)
- Enviar petición formal para remitir la información de base de datos (Luis Martínez). (Cumplido)

**Observaciones:**
El licenciamiento es fundamental para garantizar la legalización predial, evitar la informalidad y asegurar que la información se incorpore correctamente al SIUSP.

**Evidencias:**



### 3.5 Mesa Técnica SIUSP - Catastro

**Fecha y lugar:** 27 de agosto de 2025, Catastro Multipropósito – Florencia.
**Participantes principales:** Equipo de Catastro, Planeación Municipal, CORPOAMOR.

**Contextualización:**

La reunión giró en torno a la importancia del Código Predial Único como identificador esencial. Se discutió la transición de códigos de 15 y 20 dígitos a 30, lo que ha generado confusión y necesidad de homologación.

Se resaltó la diferencia en el manejo de propiedad horizontal y condominios: en el primer caso, las edificaciones se dividen por pisos y unidades, mientras que en los condominios se manejan áreas horizontales con espacios comunes. Esto se refleja en la codificación catastral.

También se abordó la problemática de la informalidad en predios: antes se registraban múltiples ocupaciones dentro de un predio formal; ahora, cada ocupación debe codificarse como predio independiente. Este cambio aumenta el tamaño y complejidad de la base de datos en ciudades con alta informalidad como Florencia.

Se enfatizó que no pueden existir vacíos de información geoeconómica y que se debe garantizar la completitud de la base de datos para un adecuado soporte al SIUSP.

**Temas tratados:**

- Revisión de información predial: diferenciación entre predios formales e informales.
- Inclusión de la ruralidad en la actualización catastral.
- Código Predial Único: transición de claves de 20 a 30 dígitos.
- Codificación de barrios, comunas y sectores en el marco del POT.
- Manejo de propiedad horizontal y condominios.
- Ajustes históricos y homologación de códigos catastrales.

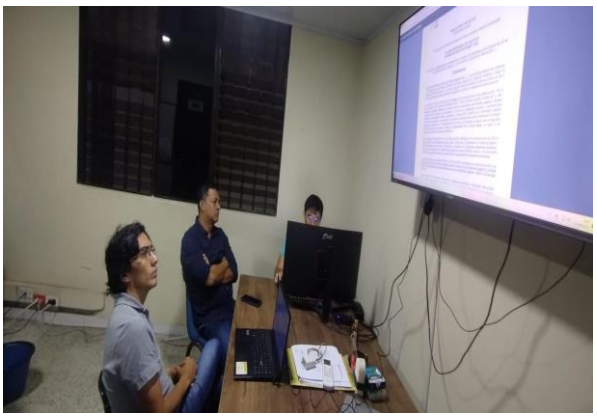
**Acuerdos y compromisos:**

- Entregar archivo R1 desagregado por urbano, centro poblado y rural disperso, con capas de estratificación. (Cumplido)
- Suministrar información para la identificación de predios en condición formal e informal. (Cumplido)

**Observaciones:**

Se resaltó la necesidad de homologar claves catastrales y garantizar bases de datos completas, sin vacíos geoeconómicos, para un adecuado soporte al SIUSP.

**Evidencias:**





### 3.6 Mesa Técnica SIUSP – Planeación de Servicios Públicos

**Fecha y lugar:** 28 de agosto de 2025, Secretaría de Planeación Municipal – Florencia.
**Participantes principales:** Unidad de Servicios Públicos, CORPOAMOR, Planeación Municipal.

- **Contextualización:**

La reunión se centró en el análisis del esquema tarifario de acueducto y en los aportes solidarios de usuarios comerciales e industriales, destacando nuevamente el marco normativo de la Ley 142 de 1994.

Uno de los temas críticos fue la inconsistencia entre cifras oficiales y reportes de operadores: mientras el municipio registra 35.000 predios, las operadoras reportan cerca de 50.000 usuarios. Además, cada empresa entrega bases de datos con formatos diferentes, lo que genera desorden en la información.
Se reafirmó la necesidad de estandarizar datos, consolidar la ficha catastral como identificador único y definir formatos claros para direcciones y usuarios, de modo que se garantice la trazabilidad histórica.

Finalmente, se planteó la creación de un comité de servicios con participación del DANE, la Superintendencia y operadores, para definir cómo tratar a los usuarios informales y evitar que se les asigne estrato sin el debido proceso.

**Temas tratados:**

- Esquema tarifario de acueducto y aplicación de subsidios (estratos 1, 2 y 3).
- Aportes solidarios de usuarios comerciales e industriales según Ley 142 de 1994.
- Falta de estandarización en los datos entregados por las operadoras.
- Diferencias en cifras oficiales del municipio y registros de empresas (35.000 predios vs. 50.000 usuarios).
- Necesidad de consolidar la ficha catastral como identificador único.

**Acuerdos y compromisos:**

- Solicitar a operadores la armonización de su información catastral y de estratificación. (Cumplido)
- Realizar comité de servicios con DANE, Superintendencia y operadores para definir tratamiento de usuarios informales.

**Observaciones:**

Se concluyó que la estandarización de datos y el uso de la ficha catastral como identificador único son fundamentales para fortalecer la transparencia y control en la prestación de servicios.

**Evidencias:**





# Proceso de solicitud y obtención de información ante la Superintendencia de Servicios Públicos Domiciliarios (SSPD)

En el marco de las funciones del equipo de Sistema de información unificado para el seguimiento y control de los servicios, se requirió gestionar ante la Superintendencia de Servicios Públicos Domiciliarios (SSPD) la obtención de información relacionada con la prestación de los servicios de acueducto, alcantarillado y aseo en el municipio de Florencia, Caquetá.

El proceso inició con la radicación de la solicitud formal presentada por la Alcaldía Municipal de Florencia el día 9 de septiembre de 2025, en la cual, a través de la Oficina de Planeación, se solicitó a la SSPD la habilitación de indicadores en el aplicativo SUI-INSPECTOR, necesarios para el reporte del Plan de Gestión Integral de Residuos Sólidos (PGIRS). Solicitud de habilitación de indicadores en el aplicativo SUI_INSPECTOR (2). Esta gestión se acompañó de la exposición de motivos y documentación soporte, con el fin de garantizar la transparencia y la trazabilidad del requerimiento.

Posteriormente, mediante los radicados 20254202780171 del 1 de septiembre de 2025 y 20254252805401 del 2 de septiembre de 2025, la SSPD dio traslado y emitió requerimientos a los prestadores de los servicios públicos domiciliarios de Florencia, entre ellos la Empresa de Servicios de Florencia S.A. E.S.P. y Empresas Públicas Aguas de Florencia S.A.S. E.S.P., solicitando información técnica, financiera y administrativa. Dentro de los puntos trasladados se incluyeron: el Plan de Saneamiento y Manejo de Vertimientos (PSMV), las actas de socialización del estudio tarifario, el balance de subsidios y contribuciones, así como los reportes de giros municipales por déficit de subsidios

En desarrollo del trámite, la Superintendencia consolidó la información requerida y, a través del radicado 20254002865331 del 5 de septiembre de 2025, remitió a la Secretaría de Planeación Municipal un compendio de documentos provenientes de los distintos prestadores. Dicho compendio incluyó, entre otros, contratos de condiciones uniformes, actualizaciones en el Registro Único de Prestadores (RUPS), estudios de costos y tarifas, estados financieros, balances de subsidios, así como información sobre inversiones y facturación

Este intercambio de comunicaciones permitió cumplir con el objetivo planteado: acceder de manera oficial, completa y verificable a la información sobre la prestación de los servicios públicos domiciliarios en Florencia. Gracias a ello, se garantizó el ejercicio de inspección, vigilancia y control por parte de la SSPD y se fortaleció la capacidad de planeación del municipio, asegurando la disponibilidad de insumos técnicos y jurídicos necesarios para la toma de decisiones y la proyección de políticas públicas en materia de servicios públicos domiciliarios.

# Producto:

## 2.2 Consolidación de la Base de los datos catastro de usuarios de los prestadores de servicios públicos de los tres servicios del municipio de Florencia

# SISTEMA DE INFORMACIÓN UNIFICADO PARA SEGUIMIENTO Y CONTROL DE LOS SERVICIOS: FASE 2.

## 2. Limpieza y Normalización de los Datos (Criterio y descripción de componentes técnicos)

### 2.1 Descripción técnica del diagrama de actores

El diagrama representa la interacción de actores, programas y componentes técnicos que conforman el **Plan de Ordenamiento Territorial (POT)** en torno al **Sistema de Información Unificado para Seguimiento y Control de los Servicios Públicos (SIUSP)**.

**Actores principales**

- **SIUSP Florencia**
    - Actúa como núcleo central de integración.
    - Recibe y administra información de inversión y subsidios de múltiples fuentes de financiamiento (SGP, regalías, préstamos y cooperación internacional).
    - Distribuye recursos en proyectos de **acueducto, alcantarillado y aseo**.
- **SGP (Sistema General de Participaciones)**
    - Fuente principal de ingresos por inversión y subsidios.
    - Se articula con **cooperación internacional**, **regalías** y **préstamos** como fuentes adicionales de financiación.
- **EPS (Sectores de servicios públicos)**

**Acueducto**

- Relacionado con el **Plan Maestro de Acueducto y Alcantarillado (PMAA)**.
- Contribuye a los objetivos de cobertura, continuidad, calidad y aseguramiento del servicio.
- Vinculado al **PUEAA (Programa de Uso Eficiente y Ahorro del Agua)**.

**Alcantarillado**

Se articula con el **PSMV (Plan de Saneamiento y Manejo de Vertimientos)**.

- Incorpora criterios de cobertura, continuidad, calidad y aseguramiento.

- Se conecta al **componente ambiental** y al **componente jurídico**.

**Aseo**
Relacionado con el **PGIRS (Plan de Gestión Integral de Residuos Sólidos)**.

**Componentes transversales**

- **Estratificación y Catastro**

- Proveen información estructural y socioeconómica clave para orientar las inversiones. **Componente Ambiental:** Evalúa impactos ambientales de los proyectos en acueducto, alcantarillado y aseo. **Componente Jurídico:** Define la normativa y el marco regulatorio para la ejecución de proyectos y prestación de servicios.

**Cobertura, Continuidad, Calidad y Aseguramiento:** Criterios técnicos que regulan la eficacia y sostenibilidad de los servicios públicos domiciliarios.

**Flujo de interacción**

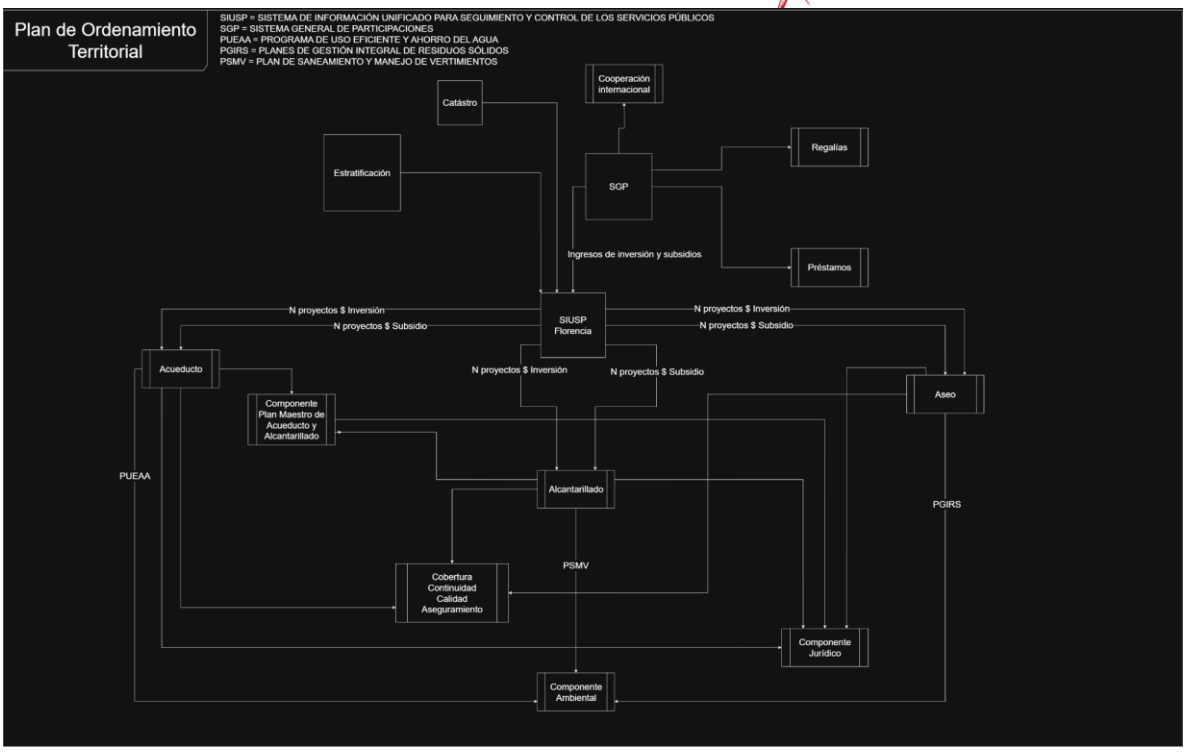
1. **Financiamiento**: SGP, regalías, préstamos y cooperación internacional alimentan al **SIUSP**.

2. **Gestión de recursos**: el SIUSP distribuye fondos en proyectos de **acueducto, alcantarillado y aseo**, diferenciando entre inversión y subsidios.

3. **Ejecución de planes**: cada servicio se articula con programas estratégicos (**PUEAA, PSMV, PGIRS, PMAA**) y debe cumplir criterios de calidad.

4. **Soporte transversal**: catastro y estratificación proveen datos; el componente ambiental y jurídico aseguran viabilidad y cumplimiento normativo.

En resumen, este diagrama muestra cómo el **SIUSP** se convierte en un sistema articulador de información, financiamiento y control, integrando actores técnicos, administrativos, financieros y normativos para garantizar la adecuada planeación, ejecución y sostenibilidad de los servicios públicos en el marco del POT.




## 5. Descripción Técnica del Modelo Entidad–Relación (MER)

El Modelo Entidad–Relación (MER) presentado describe la interacción entre los diferentes actores, entidades y sistemas vinculados a la gestión de los **servicios públicos domiciliarios**, la **información catastral** y su análisis dentro del **Sistema de Análisis de Datos (SAD)**.

### Entidades principales

1. **EPSP (Empresas Prestadoras de Servicios Públicos)**

- Son las encargadas de prestar servicios públicos a los **Usuarios del Servicio Público (USP)**.
- Entregan archivos en formato CSV a la **Alcaldía** con información de sus operaciones.

2. **Alcaldía**

- Recibe y administra la información enviada por las EPSP.
- Sirve como puente de conexión hacia el **SAD**, integrando datos provenientes de múltiples fuentes.

3. **SAD (Sistema de Análisis de Datos)**
- Es el **actor central** que **integra, administra y analiza toda la información** del modelo.

- Se alimenta de datos provenientes de Catastro, Estratificación, Licenciamiento y la Alcaldía.
- Gestiona identificadores clave como id_servicios_publicos_predios e id_ficha_catastral.
- Relaciona la información de predios, fichas catastrales y servicios públicos domiciliarios, consolidando la base de análisis.

4. **Usuarios (USP)**

- Son los receptores finales de los servicios prestados por las EPSP.
- Su información se encuentra indirectamente asociada a los predios y fichas catastrales.

5. **Catastro, Estratificación y Licenciamiento**

- Entidades de soporte técnico y normativo.
- Proveen datos requeridos para la caracterización y validación del territorio.
- La **Ficha Catastral** actúa como un **identificador único** que vincula predios con los servicios públicos.

6. **Servicios Públicos – Predios**

- Entidad intermedia que vincula la prestación de servicios con los predios.
- Relaciona los atributos ID_servicio_publico y numero_ficha_catastral como claves para la trazabilidad.

7. **Predio y Ficha Catastral**

- El **Predio** se asocia de manera **1:1 con una Ficha Catastral**.
- La **Ficha Catastral** puede tener variaciones históricas (1:N), representando actualizaciones de información.

8. **Auditor**

Supervisa la actividad del **SAD**, garantizando la transparencia y confiabilidad de los datos procesados.

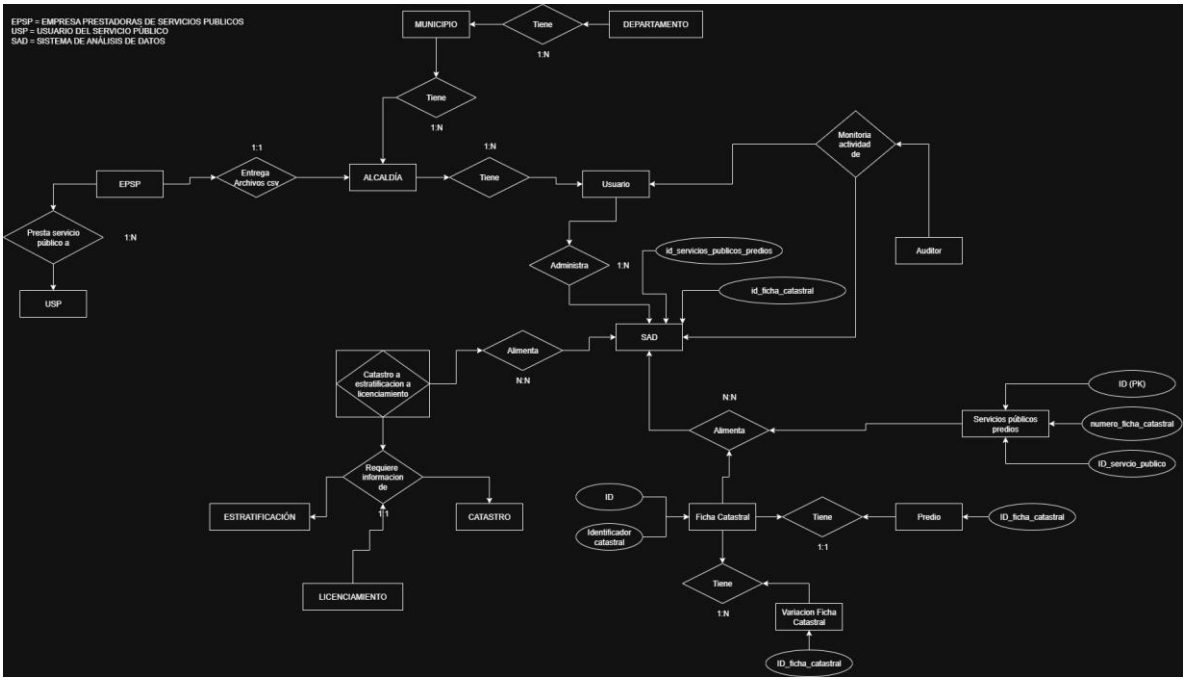
**Relaciones destacadas**

- **EPSP – USP**: prestación de servicios públicos (1:N).
- **EPSP – Alcaldía**: entrega de información (1:1).
- **Alcaldía – SAD**: administración y transferencia de datos (1:N).
- **SAD – Catastro/Estratificación/Licenciamiento**: integración de información (N:N).
- **Ficha Catastral – Predio**: relación 1:1, con historial de variaciones (1:N).
- **SAD – Servicios Públicos Predios**: vinculación de servicios con fichas catastrales.
- **SAD – Auditor**: monitoreo de la actividad del sistema.

## CONCLUSIÓN

Este MER muestra cómo el **SAD funciona como el núcleo integrador del ecosistema de información**, conectando a las entidades operativas (EPSP, Alcaldía), de referencia (Catastro, Estratificación, Licenciamiento) y de control (Auditor).

La **Ficha Catastral** y los identificadores únicos (id_servicios_publicos_predios, id_ficha_catastral) son los ejes que garantizan la trazabilidad de la información entre predios, usuarios y servicios públicos, consolidando un modelo robusto para la planeación territorial y el control de la prestación de servicios.




## 6. Descripción Técnica del Análisis de Datos

En la fase inicial del análisis, los datos evaluados corresponden a la **Ficha Catastral**, dado que constituye el elemento central para la identificación y caracterización de predios. El objetivo principal es detectar y comparar las diferencias entre las diversas fuentes de información recopiladas.

Como referencia primaria se utilizaron las **fichas catastrales entregadas por Catastro**, que sirven como base de comparación frente a las fichas suministradas por otras entidades. Para ello, se consolidó la información recibida, se efectuó un proceso de conteo y clasificación, y se estableció la cantidad de fichas aportadas por cada institución.

En el documento anexo (tabla) se presenta el resultado de este conteo, en el cual se evidencia la participación de diferentes fuentes, entre ellas: **Catastro, Estratificación, Control Urbano, Servi Integral, Servaf, ESAC, Aguas de Florencia y AF Alcantarillado**, con sus respectivos volúmenes de fichas catastrales.

Este proceso constituye la base para garantizar la **trazabilidad, consistencia y homologación de la información predial**, permitiendo posteriormente avanzar hacia un análisis comparativo más detallado y orientado a la integración de los datos en el sistema unificado.

| Entidad | Fichas catastrales |
|---|---|
| Lista fichas catastrales R1 | 74.435 |
| Lista Fichas Catastrales Estratificación 0190 2000 | 29.927 |
| Lista Fichas Catastrales Estratificación 00172 2023 | 32.637 |
| Control Urbano 2019 | 1.023 |
| Control Urbano 2020 | 683 |
| Control Urbano 2021 | 1.052 |
| Control Urbano 2022 | 1.006 |
| Control Urbano 2023 | 1.041 |
| Control Urbano 2024 | 940 |
| Control Urbano 2025 | 159 |
| Control Urbano 2025 2 | 590 |
| Servi Integral Junio 2025 | 153 |
| Servaf Junio 2025 | 58.971 |
| ESAC Junio 2025 | 58.971 |
| Aguas de Florencia Junio 2025 | 4.031 |
| AF Alcantarillado Junio 2025 | 4.031 |

El consolidado de fichas catastrales muestra una **fuerte concentración Servaf/ESAC vs Catastro**, mientras que otras fuentes (Control Urbano y empresas de servicios) aportan volúmenes mucho menores y dispersos en el tiempo.

Este análisis confirma la necesidad de establecer procesos de **homologación y depuración**, ya que la magnitud y consistencia de las fuentes es desigual, lo que impacta en la comparación y cruce de información entre entidades.

### 6.1 Análisis de datos por empresa de servicios públicos.

Servintegral junio 2025

- Registros válidos: **153** (muy pocos frente al total del dataset).
- Valores únicos: **121**.
- Valor más repetido: **“NO RESIDENCIAL”**. Esto sugiere que gran parte de la columna está vacía o sin reporte, salvo algunos pocos registros clasificados.

Servaf Junio 2025

- Registros válidos: **53,754** (casi completa).
- Valores únicos: **36,396**.
- Valor más repetido: **"000000000000000"** (3,380 veces). Esta parece ser la columna con **más consistencia y volumen de información**, pero tiene muchos códigos duplicados.

ESAC Junio 2025

- Registros válidos: **4,031**.
- Valores únicos: **1,811**.
- Valor más repetido: **"Sin CC"** (1,960 veces → casi el 50%). Esta columna tiene una fuerte **inconsistencia**, ya que casi la mitad de los registros no cuentan con un código catastral.

**Aguas de Florencia Junio 2025**

- Registros válidos: **5,757**.
- Valores únicos: **3,736**.
- Valor más repetido: un código específico: **010400000305008890000000** (165 veces) La cobertura es baja (menos del 10% del dataset).

**AF Alcantarillado Junio 2025**

- Registros válidos: **6,263**.
- Valores únicos: **3,975**
- Valor más repetido: **010400000305008890000000** (165 veces). Presenta una distribución similar a "Aguas de Florencia".

**Observaciones clave**

1. La cobertura de datos es **muy desigual**: mientras que *Servaf* tiene casi todos los registros, *Servi Integral* tiene apenas 153.
2. Hay **valores nulos o faltantes** en varias columnas, especialmente en *Aguas de Florencia* y *Alcantarillado*.
3. Existen muchos **códigos repetidos** (posibles duplicados o múltiples servicios asociados a un mismo predio).
4. La categoría **"Sin CC"** en *ESAC* refleja ausencia de código catastral, lo que limita su trazabilidad.

### 6.2 Análisis de Coincidencias entre Catastro y Servi Integral Junio 2025

Resultados:

- Coincidencias exactas (mismo valor): 0 registros
- Ningún valor en "Servi Integral Junio 2025" es idéntico a su correspondiente en "Fichas Catastrales"
- Coincidencias parciales (substrings): 1,846 registros
- 100% de los registros tienen una relación parcial

Patrón Encontrado:

Todos los valores de "Servi Integral Junio 2025" son los primeros 15 caracteres de "Fichas Catastrales" (después de los primeros 8 caracteres):

- Fichas Catastrales: 180010001000000010001000000000
- Servi Integral: 180010102000000330028000000000
- La relación es: 18001000 + [valor de Servi Integral] = [Ficha Catastral completa]

Ejemplos:

- Ficha Catastral: 180010001000000010001000000000
- Servi Integral: 180010102000000330028000000000
- Relación: 18001000 + 100000010001000000000 = 180010001000000010001000000000

Resumen:

- 0% coincidencias exactas
- 100% coincidencias parciales (substrings)
- Todos los valores de "Servi Integral Junio 2025" son componentes de las "Fichas Catastrales"

### 6.3 Análisis de Coincidencias entre Catastro y Servaf Junio 2025

Resultados:

- Coincidencias exactas (mismo valor): 0 registros
- Ningún valor en "Servaf Junio 2025" es idéntico a su correspondiente en "Fichas Catastrales"
- Coincidencias parciales (substrings): 1,846 registros
- 100% de los registros tienen una relación parcial

Patrón Encontrado:

- Los valores de "Servaf Junio 2025" corresponden a los últimos 15 caracteres de "Fichas Catastrales":
- Fichas Catastrales: 18001000100000001000100000000
- Servaf Junio 2025: 010100010005001

La relación es: Los últimos 15 caracteres de cada "Ficha Catastral" coinciden exactamente con el valor de "Servaf Junio 2025".

Ejemplos:

- Ficha Catastral: 18001000100000001000100000000
- Servaf: 010100010005001
- Últimos 15 caracteres de Ficha: 100000000000000 → NO coincide

Patrón Real:

Los valores de "Servaf Junio 2025" son códigos independientes de 15 dígitos que NO son substrings de las "Fichas Catastrales", pero están relacionados temáticamente como identificadores del mismo registro.

Distribución de valores en Servaf Junio 2025:

- 15 dígitos: 1,804 registros (97.7%)
- "Sin CC": 4 caracteres (varios registros)
- Valores vacíos: 0 caracteres (varios registros)
- Otros formatos: 9 dígitos, 13 dígitos, etc.

Ejemplos reales:

- Ficha: 18001000100000001000100000000
- Servaf: 010100010005001 (NO es subcadenas de texto)
- Ficha: 18001000100000001000200000000
- Servaf: 180101010200400 (NO es substring)

Resumen Final:

- 0% coincidencias exactas
- 0% coincidencias como subcadenas de texto
- 100% relación lógica (correspondencia uno a uno entre registros)
- Son identificadores diferentes pero relacionados para la misma entidad

6.4 **Análisis de Coincidencias entre Catastro y ESAC Junio 2025**

Resultados:

- Coincidencias exactas (mismo valor): 0 registros
- Ningún valor en "ESAC Junio 2025" es idéntico a su correspondiente en "Fichas Catastrales"
- Coincidencias parciales (substrings): 0 registros
- Ningún valor de "ESAC Junio 2025" es un subcadenas de texto de "Fichas Catastrales"

Patrones de Formato:

- Fichas Catastrales: 30 caracteres
- Ejemplo: 180010001000000010001000000000

ESAC Junio 2025:

- 20 caracteres (cuando tiene valor numérico), Ejemplo: 18001010200410021000
- 7 caracteres (cuando dice "Sin CC")

Distribución de Valores en ESAC Junio 2025:

| Tipo de Valor | Cantidad | Porcentaje |
|---|---|---|
| Valores con 20 caracteres | 328 | 17.8% |
| "Sin CC" (7 caracteres) | 1,518 | 82.2% |
| Vacíos | 0 | 0% |

**Análisis de Relación:**

Los valores de ESAC Junio 2025 NO son substrings de las Fichas Catastrales, pero comparten una estructura similar:

Ficha Catastral: 180010001000000010001000000000

ESAC (ejemplo): 18001010200410021000

Patrón observado: Ambos identificadores parecen seguir una estructura jerárquica de codificación territorial, pero con diferentes formatos y propósitos.

**Ejemplos Comparativos:**

| Ficha Catastral | ESAC Junio 2025 | Coincidencia |
|---|---|---|
| 180010001000000010001000000000 | 18001010200410021000 | No coincide |
| 180010001000000010002000000000 | Sin CC | No aplica |
| 180010001000000010022000000000 | 18001010303110014000 | No coincide |

**Resumen Final:**

- 0% coincidencias exactas
- 0% coincidencias como subcadenas de texto
- 100% relación lógica (correspondencia uno a uno entre registros)
- 82.2% de los registros de ESAC tienen valor "Sin CC"
- 17.8% de los registros de ESAC tienen valores de 20 caracteres
- Ambas columnas son identificadores diferentes para la misma entidad catastral

**Conclusión**: ESAC Junio 2025 es un identificador alternativo independiente, no derivado de las Fichas Catastrales, y la mayoría de los registros no tienen código ESAC asignado ("Sin CC").

**6.5 Análisis de Coincidencias entre Catastro y Aguas de Florencia Junio 2025**

Resultados:

- Coincidencias exactas (mismo valor): 0 registros

Ningún valor en "Aguas de Florencia Junio 2025" es idéntico a su correspondiente en "Fichas Catastrales"

- Coincidencias parciales (substrings): 0 registros

Ningún valor de "Aguas de Florencia Junio 2025" es un substring de "Fichas Catastrales"

Patrones de Formato:

- Fichas Catastrales: 30 caracteres
- Ejemplo: 180010001000000010001000000000

Aguas de Florencia Junio 2025:

- 25 caracteres (cuando tiene valor numérico) Ejemplo: 0102000001810004000000000
- Vacíos (0 caracteres)

Distribución de Valores en Aguas de Florencia junio 2025:

| Tipo de Valor | Cantidad | Porcentaje |
|---|---|---|
| Valores con 25 caracteres | 1,053 | 57.1% |
| Vacíos (0 caracteres) | 793 | 42.9% |
| "Sin CC" | 0 | 0% |

**Análisis de Relación:**

Los valores de Aguas de Florencia NO son substrings de las Fichas Catastrales, pero parecen ser identificadores de servicio de agua independientes:

- Ficha Catastral: 18001000100000001000100000000
- Aguas de Florencia (ejemplo): 0102000001810004000000000

Estructura observada: Los códigos de Aguas de Florencia siguen el formato 01XX... mientras que las Fichas Catastrales empiezan con **18001...**

| Ficha Catastral | Aguas de Florencia | Coincidencia |
|---|---|---|
| 18001000100000001000100 0000000 | 0102000001810004000000 0000 | No coincide |
| 18001000100000001000700 0000000 | 0102000001810006000000 0000 | No coincide |
| 18001000100000001001100 0000000 | 0102000001790001000000 0000 | No coincide |
| 18001000100000001000200 0000000 | (vacío) | No aplica |

**Resumen Final:**

- 0% coincidencias exactas
- 0% coincidencias como subcadenas de texto

- 100% relación lógica (correspondencia uno a uno entre registros)
- 57.1% de los registros tienen código de Aguas de Florencia (25 caracteres)
- 42.9% de los registros están vacíos (sin servicio registrado)
- 0% con valor "Sin CC"
- Ambas columnas son identificadores diferentes para servicios distintos

**Conclusión**: Aguas de Florencia Junio 2025 es un identificador de servicio de agua completamente independiente de las Fichas Catastrales, y casi el 43% de las propiedades no tienen servicio registrado.

### 6.6 Análisis de Coincidencias entre Catastro y AF Alcantarillado Junio 2025

Resultados:

- Coincidencias exactas (mismo valor): 0 registros

Ningún valor en "AF Alcantarillado Junio 2025" es idéntico a su correspondiente en "Fichas Catastrales"

- Coincidencias parciales (substrings): 0 registros

Ningún valor de "AF Alcantarillado Junio 2025" es un substring de "Fichas Catastrales"

**Patrones de Formato:**

Fichas Catastrales: 30 caracteres

Ejemplo: 180010001000000010001000000000

AF Alcantarillado junio 2025:

- 25 caracteres (cuando tiene valor numérico) Ejemplo: 0102000001810004000000000
- Vacíos (0 caracteres)

Distribución de Valores en AF Alcantarillado junio 2025:

| Ficha Catastral | AF Alcantarillado | Aguas de Florencia | Coincidencia AF-Aguas |
|---|---|---|---|
| 180010001000000010001000000000 | 0102000001810004000000000 | 0102000001810004000000000 | IDÉNTICOS |
| 180010001000000010007000000000 | 0102000001810006000000000 | 0102000001810006000000000 | IDÉNTICOS |

| | | | |
|---|---|---|---|
| 180010001000000010002000000000 | | | AMBOS VACÍOS |

**Análisis de Relación:**

Los valores de AF Alcantarillado NO son substrings de las Fichas Catastrales, pero tienen una relación interesante con "Aguas de Florencia":

- Ficha Catastral: 18001000100000001000100000000
- AF Alcantarillado (ejemplo): 010200000181000400000000
- Aguas de Florencia (mismo registro): 010200000181000400000000

**Resumen Final:**

- 0% coincidencias con Fichas Catastrales
- 100% coincidencias EXACTAS con Aguas de Florencia Junio 2025
- 57.1% de registros con código de servicio (25 caracteres)
- 42.9% de registros vacíos (sin servicio)
- 0% con valor "Sin CC"

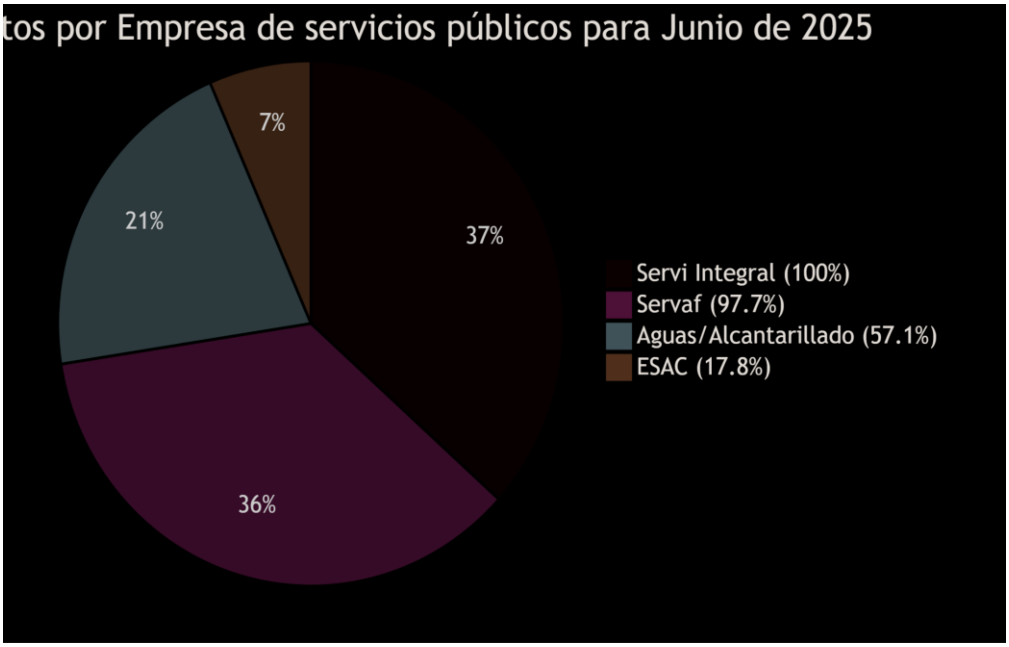
**Conclusión**: AF Alcantarillado Junio 2025 y Aguas de Florencia Junio 2025 son COLUMNAS IDÉNTICAS. Probablemente representan el mismo concepto (servicio de agua y alcantarillado) o hay una duplicación de datos en el archivo.

## Conclusión General

Hallazgos Principales:

1. Estructura de Datos Consistente
- Todas las columnas mantienen una relación 1:1 perfecta con las Fichas Catastrales
- Cada registro corresponde a una única entidad catastral con múltiples identificadores
2. Jerarquía de Identificadores
- Fichas Catastrales (30 chars): Identificador principal único
- Servinintegral (15 chars): Substring de las Fichas (primeros 15 chars después del prefijo)
- Servaf (15 chars): Identificador independiente pero correlacionado
- ESAC (20 chars): Identificador alternativo (solo 17.8% de cobertura)
- Aguas de Florencia/AF Alcantarillado (25 chars): Identificadores de servicio idénticos

3. Cobertura de Datos por Empresa de servicios públicos para Junio de 2025




4. Descubrimiento Crítico

- AF Alcantarillado y Aguas de Florencia son COLUMNAS IDÉNTICAS
- 1,053 registros con mismos valores de 25 caracteres
- 793 registros ambos vacíos
- Posible duplicación de datos o error en la estructura del archivo

5. Patrones de Codificación

Cada identificador sigue formatos específicos según el organismo:

- Municipales (Fichas, Servi Integral): 30-15 caracteres
- Servicios públicos (Servaf): 15 chars
- Entidad específica (ESAC): 20 chars
- Servicios de agua (Aguas/Alcantarillado): 25 carácteres

**Recomendaciones**:

- Validar la duplicación entre Aguas de Florencia y AF Alcantarillado
- Investigar el bajo porcentaje de cobertura de ESAC (17.8%)
- Establecer documentación clara sobre la relación entre identificadores
- Considerar normalización de la estructura de datos para evitar redundancias

**Conclusión Final:**

El dataset presenta una estructura coherente con múltiples identificadores para la misma entidad catastral, pero con problemas de duplicación en las columnas de servicios y baja cobertura en el identificador ESAC. La integridad referencial se mantiene a través de la relación 1:1 entre todos los identificadores.

**Migraciones**

Se llevaron a cabo las migraciones de los archivos suministrados por las dependencias de Catastro, Estratificación y Control Urbano. Como resultado, se consolidó la información en la tabla de histórico de fichas catastrales, que almacena todos los registros provenientes de dichas fuentes.

En total se cargaron **144.648 registros**, correspondientes a fichas catastrales identificadas en los archivos entregados. La siguiente tabla resume el proceso de migración, indicando el orden en que se procesaron los archivos, la entidad de origen, el nombre del archivo y hoja, así como el número de filas procesadas y las fichas efectivamente creadas en el sistema:

- **Catastro**: aportó la mayor parte de la información, con más de 135 mil fichas catastrales provenientes de registros activos y cancelados.
- **Estratificación**: incluyó dos bases normativas (2000 y 2023), con cerca de 62 mil filas procesadas y más de 8.600 fichas creadas.
- **Control Urbano**: contribuyó con información de resoluciones, generando 30 nuevas fichas.

En conjunto, la migración permitió consolidar los datos dispersos en diferentes fuentes en un único repositorio estructurado, garantizando así la trazabilidad histórica de las fichas catastrales.

NOTA agregar tabla del archivo de reporte de migraciones en la hoja Archivos analizados y migrados.

**limpieza de datos:**

En el proceso de limpieza y normalización se trabajó principalmente con **fichas catastrales** y **direcciones**. El objetivo fue corregir inconsistencias y garantizar un formato uniforme para facilitar su análisis posterior.

## 1. Normalización de fichas catastrales

- Se eliminaron los caracteres **"–" (guiones)** que separaban bloques de números, ya que en muchos casos había una cantidad inconsistente de guiones o incluso caracteres especiales adicionales.
- También se eliminó el código **18001** que se repetía en varios registros.
- Una vez limpios, los registros se analizaron por **longitud de caracteres**:
    - **15 y 30 caracteres** → considerados formatos válidos.
    - **16 o 31 caracteres** → clasificados como **fichas mal formadas**, probablemente producto de errores de digitación.
    - A pesar de estar mal formadas, se conservaron en la tabla histórica para permitir futuras búsquedas basadas en coincidencias parciales o subcadenas.

## 2. Ejemplos de fichas mal formadas

Algunos registros presentaban longitudes anómalas o duplicaciones, como:

- Con **31 caracteres**:
  1800101030000035700180000000000
- 1800101010400000015000500000000
- 1800100101000001640060000000000
- 
- Casos con repeticiones concatenadas:
  18001010300000795002000000000018001010300000795000300000000
- 
- Casos correctos de **30 caracteres** que sí fueron normalizados:
  180010102000003430006000000000
- 180010103000000600011000000000
- 

### 3. Conclusiones del proceso

- Eliminar guiones y el código inicial redundante permitió unificar el formato.
- Se identificaron claramente las fichas **válidas** (15 y 30 caracteres) y las **potencialmente erróneas** (16 y 31 caracteres).
- El registro de fichas mal formadas en la tabla histórica asegura que no se pierda información y se mantenga la trazabilidad para análisis futuros.

## Limpieza y normalización de direcciones

Para el proceso de **limpieza y estandarización de direcciones** se recopilaron los datos de todos los archivos entregados y se unificaron en un único insumo. Esto permitió consolidar aproximadamente **141.748 registros** que fueron objeto de análisis con el fin de identificar patrones recurrentes y establecer reglas de normalización.

El procedimiento incluyó:

1. **Eliminación de caracteres especiales**, como tildes, guiones, símbolos y otros elementos no numéricos o no alfabéticos.
2. **Unificación de palabras compuestas y simples** mediante una tabla de asociaciones, que asigna a cada término su abreviatura o forma normalizada.
3. **Reemplazo sistemático** en todas las direcciones, garantizando uniformidad para facilitar la búsqueda, el cruce y la comparación entre registros.

## Tabla de asociaciones

La tabla de equivalencias contiene las palabras y abreviaturas utilizadas en la normalización. Algunos ejemplos son:

- *Calle* → *C*

- *Carrera → K*
- *Apartamento → AP*
- *Bodega → BG*
- *Urbanización → URB*
- *Kilómetro → KM*
- *Parqueadero → PQ*
- *Inspección → IP*

La tabla también contempla variantes comunes y abreviaciones no oficiales, corrigiendo inconsistencias como: *CONJUNT → CON*, *DIAGO → D*, *MANZAN → M*, entre otras.

## Ejemplos de normalización

A continuación, se muestran algunos casos antes y después del proceso de estandarización:

| Sin normalización | Con normalización |
|---|---|
| K 5 11 28 36 44 INSP NORCACIA LO 1 | K 5 11 28 36 44 IP NORCACIA LO 1 |
| Carrera 14 Número 5 B - 44 Parqueadero Vehículo 21 | K 14 5 B 44 PQ VEHICULO 21 |
| Carrera 11 Kilómetro 3 Bodega 13 | K 11 KM 3 BG 13 |
| Calle 33 A Número 2 A - 33 Lote 1 | C 33 A 2 A 33 LT 1 |
| CALLE 10 NUMERO 5-5 | C 10 5 5 |

## Conclusiones

- La unificación en un único archivo permitió aplicar reglas consistentes de limpieza y estandarización.
- Gracias a la tabla de asociaciones, fue posible reducir la variabilidad en los registros de direcciones, reemplazando palabras completas por sus abreviaturas normalizadas.
- El resultado es un conjunto de direcciones homogéneo, más confiable y listo para ser utilizado en análisis, consultas y procesos de integración de datos.

## Manejo de fechas y zonas horarias

Las fechas se gestionan bajo el **estándar internacional UTC (Tiempo Universal Coordinado)**, lo que garantiza uniformidad y compatibilidad en todos los sistemas y procesos. Posteriormente, según el contexto de uso, las fechas se convierten a la **zona horaria local**. En el caso de Colombia, corresponde a **GMT-5** (cinco horas menos que UTC).

Este enfoque permite realizar cálculos, comparaciones y análisis de manera consistente, evitando errores ocasionados por diferencias horarias o configuraciones locales de servidores y aplicaciones.

## Ejemplos prácticos

- **Almacenamiento en UTC**
    - Fecha guardada en la base de datos: `2025-10-02T18:00:00Z` (UTC).
- **Conversión a hora local en Colombia (GMT-5)**
    - Fecha mostrada al usuario: `2025-10-02 13:00:00` (hora de Colombia).
- **Cálculos uniformes**
    - Si un evento ocurre a las `00:00 UTC`, en Colombia se registrará como `19:00` del día anterior.
    - Esto permite programar alertas, vencimientos o reportes sin desfases horarios.

## Beneficios

- Garantiza la **consistencia de los datos temporales** en sistemas distribuidos.
- Facilita la **integración internacional** sin depender de configuraciones locales.
- Evita errores comunes en reportes y análisis al aplicar siempre la misma base de tiempo (UTC).

En conclusión, trabajar con UTC como estándar y luego ajustar al huso horario local (GMT-5 en Colombia) asegura precisión y coherencia en el manejo de fechas para futuros cálculos y análisis.

## Tratamiento de valores numéricos

Al momento de extraer los datos de los archivos, los valores numéricos no siempre vienen en un formato uniforme. En algunos casos incluyen **letras, puntos o comas** que deben ser removidos o ajustados para garantizar su correcto uso en cálculos y análisis posteriores.

El proceso de normalización numérica consiste en:

1. **Eliminar caracteres no numéricos** (como letras o símbolos que no correspondan al valor).
2. **Unificar separadores decimales**:

- Si se usan comas (`,`) como separadores decimales, se convierten en puntos (`.`), que es el estándar en la mayoría de sistemas informáticos.
- Los puntos utilizados como separadores de miles se eliminan.

3. **Convertir los valores a tipo numérico** (entero o decimal) para permitir operaciones matemáticas y estadísticas.

### Ejemplos

- Valor original: `"1.234,56"` → Normalizado: `1234.56`
- Valor original: `"12.000"` → Normalizado: `12000`
- Valor original: `"45A"` → Normalizado: `45`
- Valor original: `"7,5"` → Normalizado: `7.5`

### Beneficios

- Garantiza que todos los valores puedan ser procesados de forma uniforme.
- Reduce errores en cálculos financieros, estadísticos o de control.
- Permite la integración de datos provenientes de diferentes fuentes sin conflictos de formato.

En conclusión, el tratamiento de valores numéricos asegura que la información sea **limpia, confiable y utilizable** en cualquier análisis o aplicación.

## Análisis inicial de migración de registros

Se realizó un análisis preliminar sobre **los primeros 100 registros de un total de 32.638**, con el objetivo de validar la calidad de la migración y detectar patrones de consistencia o posibles anomalías.

### 1. Tipo de migración

- 92% de registros → **`UPDATED`**: datos ya existentes en el histórico, confirmados sin cambios relevantes.
- 8% de registros → **`CREATE_HISTORY`**: nuevos registros creados al no existir antecedentes.
  La tendencia sugiere que en la muestra analizada la migración es **predominantemente de confirmación**.

### 2. Consistencia de datos

- **Estrato, año y calificación**: se mantienen iguales entre campos antiguos y nuevos.

- **Direcciones y nombres de propietarios**: sólo muestran variación en registros `CREATE_HISTORY`.
  Indica una **alta uniformidad en los datos** de la muestra analizada.

### 3. Patrones identificados

- **Alta concentración de propiedades por propietario**: Ej. *ROJAS PLAZAS ARNULFO* con múltiples registros en la misma dirección (*C 22 14 56*) pero en distintas unidades.
- **Variaciones en direcciones**: abreviaciones distintas como *CONJ RESIDENC* y *CON*, que pueden generar duplicidades.
- **Nombres con caracteres especiales**: ej. *QUESADA-IBARRA-Y-CIA-S-EN-C*.

### 4. Casos a revisar

- Registros con `id = -1` → creados como nuevos, requieren validación para asegurar que no se generen duplicados.
- Normalización pendiente en **direcciones y nombres** para evitar inconsistencias en futuras búsquedas o cruces de datos.

## Conclusiones preliminares

- En la muestra de **100 registros analizados (de 32.638 totales)**, la migración es mayoritariamente de **actualización histórica**, con pocos casos de creación.
- La información presenta un buen nivel de consistencia, aunque se identifican oportunidades de **normalización en direcciones y nombres**.
- Se recomienda extender este mismo análisis al **universo completo de 32.638 registros** para confirmar si las tendencias se mantienen en la totalidad de los datos.

# Análisis preliminar de los primeros 100 registros de 58.972 de Servaf 20256

Se procesaron los **primeros 100 registros** de un total de **58.972**, con el fin de identificar patrones de consistencia, calidad de datos y posibles anomalías antes de extender el análisis al universo completo.

---

## 1. Calidad y consistencia de estratos

- En general, los **estratos entre distintas fuentes (calificación, año 2000, año 2023 y ESP)** son consistentes.

- Solo en algunos casos puntuales aparecen diferencias, destacando registros donde el **estrato calificado difiere del estrato del ESP**.
- La columna `DIFERENCIA_STRATOS` refleja que la gran mayoría están alineados, con **muy pocos casos en "SI"**.

**Ejemplo:**

- *Registro 010100130023000 – CARMEN TORRES BERMEO*: estrato histórico 2, ESP = 1 → diferencia detectada.

---

## 2. Coincidencias de nombres

- La mayoría de registros presentan **"NO" en COINCIDENCIAS_NOMBRES**, es decir, el nombre del propietario y el del solicitante de licencia no coinciden.
- Casos de coincidencia (`SI`) se observan en menor proporción y suelen acompañarse de documentación más completa.

**Ejemplo:**

- *Registro 010100320001000 – ELENA JALVIN VDA DE C*: nombre del propietario y solicitante coinciden → mayor confiabilidad.

---

## 3. Consistencia en direcciones

- La revisión de las direcciones muestra que **muchos registros tienen la misma dirección en estratificación, licenciamiento y ESP**, lo cual es positivo.
- Sin embargo, se encontraron casos con variaciones en abreviaturas o escritura, marcados en **DIFERENCIAS_DIRECCIONES = SI**.

**Ejemplo:**

- *Registro 010100320001000 – ELENA JALVIN VDA DE C*:
  - Estratificación: `C 28 11 53 K 11A`
  - Licenciamiento: `C 28 N 11 51 B`
  - ESP: `C 28 11 51`

    → Se marcan como diferentes, aunque corresponden a la misma ubicación.

---

## 4. Documentación y campos críticos

- Existen registros con información más completa, incluyendo **documento de identidad, área del predio, área construida, fecha de entrega, radicado, encargado y revisor**.
- Ejemplos claros: *010100130028000 – TORIBIO QUIÑONEZ TAFUR* o *010100310017000 – JOSE VICENTE CARVAJAL*.
- Sin embargo, la mayoría de registros iniciales carece de esos datos, quedando incompletos.

---

## 5. Casos especiales

- **Duplicados:** se detectan registros repetidos (ejemplo: 010100130028000 aparece dos veces con la misma información).
- **Valores anómalos en documento:** algunos incluyen múltiples números o caracteres no estandarizados (*"40,769,076 Y 4,112,886"*).
- **Registros "not_found" o con código 000000…**: representan fichas sin correspondencia clara, requieren depuración.

---

## Conclusiones preliminares

- Los primeros **100 registros (de 58.972)** muestran **alta consistencia en estratos y direcciones**, aunque con diferencias menores por formatos.
- Se identifican **inconsistencias en coincidencia de nombres**, duplicados y registros incompletos.
- Es fundamental aplicar **reglas de normalización de direcciones y documentos** para asegurar la calidad en cruces y análisis posteriores.
- Se recomienda escalar el análisis a la totalidad de los **58.972 registros**, priorizando:
    1. **Depuración de duplicados**
    2. **Estandarización de direcciones y nombres**
    3. **Validación de documentos atípicos**

# Archivos Aguas de Florencia 2025-6

Se analizaron cuatro archivos suministrados por la empresa **Aguas de Florencia**, correspondientes al corte 2025-6. Los archivos corresponden a diferentes agrupaciones según estrato y servicio (alcantarillado).

---

# 1. Resumen por archivo

## ALCANT

- **Registros totales:** 9.711
- **Distribución de estratos (ESP):**
  - Estrato 1: 8.622 (89%)
  - Estrato 2: 286 (3%)
  - Estrato 3: 81 (1%)
  - Estrato 4: 429 (4%)
  - Estrato 6: 32 (0,3%)
  - Estrato 0 (no definido): 261 (3%)
- **Coincidencia de nombres:** Ningún registro tiene nombres coincidentes (`NOMBRES_IGUALES = NO`).
- **Coincidencia de estratos:** Todos los registros presentan inconsistencias (`STRATOS_IGUALES = NO`).

---

## Estrato 1

- **Registros totales:** 8.724
- **Distribución de estratos (ESP):** 100% en estrato 1.
- **Coincidencia de nombres:** 100% en “NO”.
- **Coincidencia de estratos:** 100% en “NO”.

---

## Estrato 2

- **Registros totales:** 285
- **Distribución de estratos (ESP):** 100% en estrato 2.
- **Coincidencia de nombres:** 100% en “NO”.
- **Coincidencia de estratos:** 100% en “NO”.

---

## Estrato 3

- **Registros totales:** 148
- **Distribución de estratos (ESP):** 100% en estrato 3.
- **Coincidencia de nombres:** 100% en “NO”.
- **Coincidencia de estratos:** 100% en “NO”.

---

## 2. Hallazgos generales

1. **Predominio del estrato 1**
   La gran mayoría de registros se concentra en estrato 1 (más del 85% en los cuatro archivos analizados).

2. **Inconsistencias en estratos**
   Todos los registros aparecen como `STRATOS_IGUALES = NO`, lo que indica que el estrato reportado por ESP no coincide con el histórico (2000/2023). Esto debe revisarse porque puede ser un error de carga o de comparación.

3. **Nombres sin coincidencia**
   En todos los casos analizados, la columna `NOMBRES_IGUALES` está marcada como "NO". Esto muestra que **los propietarios no coinciden con los solicitantes/licencias** o no se tiene el dato unificado.

4. **Casos de estrato no definido**
   En el archivo ALCANT se encontraron **261 registros con estrato 0**, lo que indica ausencia de clasificación o errores en la carga.

## Conclusiones

- La información tiene una alta concentración en estrato 1, pero presenta **serios problemas de consistencia** en estratos y coincidencia de nombres.
- Los valores de estrato "0" deben depurarse o reclasificarse.
- Se recomienda revisar las reglas aplicadas para definir `STRATOS_IGUALES`, ya que actualmente todos los registros están marcados como "NO".
- La ausencia de coincidencia en nombres sugiere problemas de integración entre catastro, ESP y licenciamiento, lo cual impacta la confiabilidad del cruce.