

GEOCODIFICADOR DE EVENTOS GEORREFERENCIADOS A NIVEL DE DIRECCIONES  
URBANAS PARA LA IDENTIFICACION DE PATRONES DELICTIVOS FATALES Y NO  
FATALES EN ZONAS ESPECÍFICAS DEL MUNICIPIO DE PASTO.

NELSON ENRIQUE QUEMÁ TAIMBUD



UNIVERSIDAD DE NARIÑO  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE SISTEMAS  
PROGRAMA DE INGENIERÍA DE SISTEMAS  
SAN JUAN DE PASTO

2016

GEOCODIFICADOR DE EVENTOS GEORREFERENCIADOS A NIVEL DE DIRECCIONES  
URBANAS PARA LA IDENTIFICACION DE PATRONES DELICTIVOS FATALES Y NO  
FATALES EN ZONAS ESPECÍFICAS DEL MUNICIPIO DE PASTO.

NELSON ENRIQUE QUEMÁ TAIMBUD

Trabajo de grado presentado como requisito parcial para optar al título de ingeniero de sistemas

Director: Ricardo Timarán Pereira, PhD.

UNIVERSIDAD DE NARIÑO  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE SISTEMAS  
PROGRAMA DE INGENIERÍA DE SISTEMAS  
SAN JUAN DE PASTO

2016

## **NOTA DE RESPONSABILIDAD**

“Las ideas y conclusiones aportadas en la tesis de grado, son responsabilidad exclusiva de sus autores”.

Artículo 1º del acuerdo No 324 del 11 de octubre de 1966, emanado del Honorable Consejo Directivo de la Universidad de Nariño.

“La Universidad de Nariño no se hace responsable de las opiniones o resultados obtenidos en el presente trabajo y para su publicación priman las normas sobre el derecho de autor”

Artículo 13, Acuerdo N. 005 de 2010 emanado del Honorable Consejo Académico.

Nota de aceptación:

---

---

---

---

---

Firma Presidente del Jurado

---

Firma del Jurado



---

Firma del Jurado



San Juan de Pasto, 2016



**You are free:**

-  to **Share** – to copy, distribute and transmit this work
-  to **Remix** – to adapt this work

**Under the following conditions:**

-  **Attribution** – You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work)
-  **Noncommercial** – You may not use this work for commercial purposes.

Subject to conditions outlined in the license.

This work is licensed under the *Creative Commons Attribution-NonCommercial 3.0 Unported* License. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc/3.0/>

or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

*Dedicado a:*

*A mi madre Aura Elisa,  
por el todo el apoyo brindado,  
por los consejos que me ha dado,  
por corregir mis errores,  
por la motivación de luchar por mis sueños,  
por orientar mi camino haciendo posible este logro,  
por su gran amor incondicional.*

*A mi padre Jesús Artemio,  
por todo el cariño y apoyo brindados,  
en el transcurso de mi vida y carrera profesional.*

*A mis hermanos William Andrés y Yesenia Elizabeth,  
por estar presentes conmigo brindándome su apoyo incondicional.*

*A mis compañeros y amigos,  
por su compañía y apoyo en los momentos difíciles.*

## **AGRADECIMIENTOS**

Al profesor Ricardo Timarán Pereira, por haberme abierto las puertas en el grupo de investigación GRIAS y motivarme a iniciar mi vida investigativa.

Al profesor Andrés Calderón Romero, por compartir conmigo el conocimiento fundamental para realizar esta investigación.

A todos los profesores, por todas sus enseñanzas y apoyo a lo largo del pregrado.

A la Gobernación de Nariño y a la Fundación Ceiba, por haber apoyado el desarrollo de esta investigación dentro del programa 'Formación de Capital Humano de Alto Nivel en el Departamento de Nariño Componente: Jóvenes Investigadores e Innovadores'.

A toda la comunidad de Software Libre, por enseñarme el valor de compartir.

## **RESUMEN**

El presente trabajo muestra la construcción del geocodificador de código abierto para el municipio de Pasto que permite traducir direcciones urbanas a coordenadas geográficas longitud/latitud. El geocodificador fue acoplado con PASTO VIEW, un visor cartográfico que permite búsqueda y ubicación de direcciones urbanas. Por otra parte, fue integrado a SIGEODEP, el sistema georreferenciado del Observatorio del Delito del Municipio de Pasto mediante la construcción del módulo de información geográfico SIGEODEP-SIG. Este nuevo módulo de SIGEODEP georreferenciará las lesiones fatales y no fatales a nivel de direcciones urbanas permitiendo visualizar zonas en el municipio donde se identifiquen situaciones atípicas o especiales sin depender de barrios o comunas como está actualmente configurado. Esto facilitará a los organismos gubernamentales y de seguridad la toma de decisiones eficaces relacionadas con la seguridad ciudadana y la prevención de delitos.

## **ABSTRACT**

The present work shows the construction of open source geocoder for the municipality of Pasto, which translates urban addresses to geographic coordinates longitude/latitude. The geocoder was coupled with PASTO VIEW, a cartographic viewer which allows the searching and localization of urban addresses. Furthermore, it was integrated to SIGEODEP, the georeferenced system of Crime Observatory of the municipality of Pasto through the construction of geographic information module denominated SIGEODEP-SIG. The new information module of SIGEODEP will enable to georeference the fatal and non-fatal injuries to urban addresses level allowing visualize areas in the municipality where atypical or special situations will be identified without depend on neighborhoods or communes as it is currently configured. Because of this, it will facilitate the effective decision making in relation to the implementation of crime prevention plans and public safety by the government and security agencies.

## CONTENIDO

INTRODUCCIÓN	16
1. MARCO TEÓRICO	21
1.1. ANTECEDENTES	21
1.2. SUPUESTOS TEÓRICOS	22
1.2.1. Información georreferenciada	22
1.2.2. Bases de datos espaciales	23
1.2.3. Modelos de representación geográfica	24
1.2.4. Sistemas de referencia espacial	26
1.2.5. Geocodificación de eventos	29
1.2.6. Visualización de información geográfica	31
1.2.7. Direcciones urbanas	33
1.2.8. Caso de estudio Observatorio del delito del municipio de Pasto	39
2. METODOLOGÍA	41
2.1. TIPO DE INVESTIGACIÓN	41
2.2. DISEÑO DE LA INVESTIGACIÓN	41
2.3. DESARROLLO DE LA INVESTIGACIÓN	42
2.3.1. Construcción del repositorio de información urbana	42
2.3.2. Construcción del geocodificador de direcciones urbanas	48
2.3.3. Construcción visor cartográfico de eventos delictivos	63
2.3.4. Pruebas de funcionamiento	73
2.3.5. Identificación y visualización geográfica de patrones delictivos	79
3. RESULTADOS DE LA INVESTIGACIÓN	87
4. DISCUSIÓN DE RESULTADOS	89
5. CONCLUSIONES	91
6. RECOMENDACIONES	94
7. BIBLIOGRAFÍA	96

## ÍNDICE DE FIGURAS

Figura 1. Georreferenciación de eventos delictivos a nivel de barrios, funcionalidad implementada en SIGEODEP. ....	18
Figura 2. Gestión de información geográfica mediante capas temáticas .....	24
Figura 3. Representación objetos geográficos: línea y polígonos (áreas). ....	25
Figura 4. Representación de la realidad en los formatos Vectorial y Raster. ....	26
Figura 5. Valores de longitud y latitud de un punto sobre la esfera. ....	27
Figura 6. Elementos de los sistemas de coordenadas geográficas: Paralelos (a), Meridianos (b) Origen de coordenadas (c).....	27
Figura 7. La superficie terrestre representada según distintas proyecciones.....	28
Figura 8. Geocodificación de eventos y despliegue de datos temáticos o características. ....	29
Figura 9. Proceso general de geocodificación .....	30
Figura 10. Elementos del área urbana que requieren asignación de nomenclaturas .....	34
Figura 11. Asignación de direcciones en el área urbana del municipio de pasto .....	35
Figura 12. Organización de elementos de direcciones basadas en la malla vial. ....	37
Figura 13. Organización de elementos de direcciones basadas en la estructura Barrio – Manzana - Predio. ....	38
Figura 14. Metodología aplicada en la investigación.....	42
Figura 15. Lado derecho: Mapas de referencia descargados del servidor Open Street Maps. Lado izquierdo: Malla vial del municipio de Pasto filtrada de los datos descargados. ....	43
Figura 16. Capas que corresponden a los barrios (lado izquierdo) y las comunas (lado derecho) del municipio de Pasto .....	44
Figura 17. Segmentación de las comunas para la recopilación de direcciones. Ejemplo: Comuna 10. ....	45
Figura 18. Address Interpolation: Complemento de la herramienta Java Open Street Maps utilizado para la recopilación de direcciones urbanas.....	46
Figura 19. Modelo Entidad - Relación de la base de datos con información del área urbana del municipio de Pasto. ....	47
Figura 20. Procesos de geocodificación para direcciones urbanas del municipio de Pasto .....	49
Figura 21. Algunas direcciones urbanas registradas por el observatorio del delito. ....	52

Figura 22. Autómata Finito Determinista que reconoce los componentes de las direcciones asignadas según la malla vial del municipio de Pasto .....	54
Figura 23. Autómata Finito Determinista que reconoce los componentes de direcciones asignadas según la nomenclatura Barrio-Manzana-Identificación de predio .....	55
Figura 24. Ejecución del procedimiento de estandarización de direcciones urbanas asignadas según la malla vial.....	56
Figura 25. Ejecución del procedimiento almacenado de estandarización de direcciones urbanas asignadas según la nomenclatura barrio - manzana - predio .....	57
Figura 26. Ejecución del procedimiento almacenado de búsqueda de direcciones urbanas asignadas según la malla vial .....	57
Figura 27. Ejecución del procedimiento almacenado de búsqueda de direcciones urbanas asignadas según la nomenclatura barrio - manzana - predio.....	58
Figura 28. Logo visor cartográfico PASTO—VIEW .....	58
Figura 29. Arquitectura del visor cartográfico PASTO-VIEW .....	59
Figura 30. Búsqueda de direcciones, funcionalidad implementada en PASTO VIEW .....	60
Figura 31. Procesamiento de archivo CSV, datos agrupados por el atributo Cluster. Funcionalidad implementada en PASTO VIEW .....	61
Figura 32. Paquete geocoder visor PASTO VIEW .....	61
Figura 33. Código de conexión del visor PASTO VIEW con el geocodificador de direcciones urbanas.....	62
Figura 34. Tablas base de datos SIGEODEP que almacenan los eventos delictivos. ....	63
Figura 35. Algunos ejemplos de datos que no corresponden a direcciones urbanas, registrados en el campo injury_id. ....	64
Figura 36. Logo SIGEODEP SIG.....	65
Figura 37. Arquitectura SIGEODEP SIG.....	65
Figura 38. Procesamiento y visualización de eventos almacenados en un archivo CSV a nivel de direcciones urbanas. ....	67
Figura 39. Visualización de mapas de puntos procesados a nivel de direcciones urbanas .....	68
Figura 40. Visualización de mapas de calor procesados a nivel de direcciones urbanas .....	68
Figura 41. Paquete Application Control .....	69
Figura 42. Paquete Geocoder.....	69
Figura 43. Clases del paquete POJO .....	71
Figura 44. Clases del Paquete DAO .....	72
Figura 45. Clases del Paquete JDBC .....	72

Figura 46. Resultados Caso de prueba 1 para el procedimiento almacenado de estandarización de direcciones urbanas .....	75
Figura 47. Resultado caso de prueba 2 Procedimiento almacenado de estandarización .....	77
Figura 48. Resultado geocodificación de direcciones de prueba por comuna.....	78
Figura 49. Resultado geocodificación de direcciones urbanas según la nomenclatura Barrio Manzana Predio.....	79
Figura 50. Visualización geográfica de los eventos delictivos que fueron agrupados al cluster0 por el algoritmo k-means.....	83
Figura 51. Visualización geográfica de los eventos delictivos que fueron agrupados al cluster1 por el algoritmo k-means.....	84
Figura 52. Visualización geográfica de los eventos delictivos que fueron agrupados al cluster2 por el algoritmo k-means.....	85
Figura 53. Visualización geográfica de los eventos delictivos que fueron agrupados al cluster3 por el algoritmo k-means.....	86

## ÍNDICE DE TABLAS

Tabla 1. Estándar de las principales abreviaturas utilizadas en el municipio de Pasto. ....	36
Tabla 2. Descripción de elementos que componen una dirección basada en la malla vial. ....	37
Tabla 3. Descripción de elementos de direcciones basadas en la estructura Barrio, Manzana Identificación de Predio.....	39
Tabla 4. Tipos de eventos delictivos asociados a las Líneas de Vigilancia del Observatorio del delito del municipio de Pasto .....	40
Tabla 5. Casos de reemplazo de símbolos y nomenclaturas en direcciones urbanas ingresadas. ....	53
Tabla 6. Resultado del proceso ejecutado por el Autómata Finito Determinista para direcciones urbanas asignadas según la malla vial.....	54
Tabla 7. Resultado del proceso ejecutado por el Autómata Finito Determinista para direcciones urbanas asignadas según la nomenclatura barrio – manzana – predio.....	55
Tabla 8. Resultados individuales caso de prueba para el procedimiento almacenado de Estandarización de direcciones urbanas. ....	74
Tabla 9. Resultados individuales Caso de Prueba 2 .....	75
Tabla 10. Resultados de geocodificación de direcciones ejecutado por comuna: direcciones asignadas según la malla vial. ....	78
Tabla 11. Resultados de geocodificación de direcciones asignadas según la nomenclatura Barrio Manzana Predio ejecutado por comuna. ....	79
Tabla 12. Atributos de los eventos delictivos no fatales seleccionados .....	80
Tabla 13. Clusters encontrados con la ejecución del algoritmo k-means.....	81
Tabla 14. Algunos de los datos procesados por el algoritmo k-means en Weka .....	82

## LISTA DE ANEXOS

ANEXO A. Manual de digitalización de direcciones urbanas del municipio de Pasto	100
ANEXO B. Instalación del geocodificador de direcciones urbanas.	113
ANEXO C. Manual de usuario visor cartográfico PASTO VIEW	119
ANEXO D. Análisis UML visor cartográfico PASTO VIEW	130
ANEXO E. Procedimientos SQL de geocodificación de eventos delictivos	135
ANEXO F. Manual de usuario visor cartográfico SIGEODEP SIG	139
ANEXO G. Análisis UML visor cartográfico SIGEODEP SIG	167

## INTRODUCCIÓN

Aproximadamente un 70% de la información que se maneja en cualquier tipo de disciplina esta georreferenciada, es decir, que hace referencia a algún evento, fenómeno o problema ubicado dentro de una zona geográfica extensa (ej. países, regiones, departamentos, comunas o barrios) o reducida (cuadras, direcciones urbanas, coordenadas geográficas longitud y latitud). Según lo mencionado por [1], los Sistemas de Información Geográfica (SIG) surgen como herramientas de apoyo en la toma de decisiones que permiten procesar esta información desde su dimensión espacial y suministrar datos para la formulación de soluciones a problemas de orden geográfico.

La georreferenciación como principal herramienta que proveen los SIG, permite el análisis de datos basado en ubicaciones específicas reconocibles de datos dentro de un área geográfica de estudio [2]. En nuestro país se han implementado sistemas de georreferenciación basados en límites geográficos específicos (regiones, departamentos, barrios, sectores, veredas, entre otros) para procesar los eventos y desplegarlos posteriormente en mapas temáticos. En muchos casos la información procesada contiene su dirección urbana de ocurrencia, una característica que de ser procesada, permitiría el análisis geográfico de los datos siniestrados con un grado alto de detalle y exactitud, desafortunadamente estos datos no se toman en cuenta debido a la ausencia de herramientas que permitan este tipo de procesamiento. El desarrollo de esta investigación permitió procesar esos datos implementando un geocodificador de eventos georreferenciados a nivel de direcciones urbanas de código abierto para el municipio de Pasto, que a través de su integración con un visor cartográfico permitan visualizar e identificar patrones visuales atípicos a partir de la ubicación geográfica de dichos eventos.

Se tuvo como caso de estudio SIGODEP, un sistema georreferenciado del observatorio del delito del municipio de Pasto, el cual cuenta con un módulo de georreferenciación a nivel de barrios y comunas. El geocodificador permitirá georreferenciar los eventos delictivos fatales y no fatales registrados por esta entidad durante los últimos 14 años y analizar el comportamiento de los datos visualizados a nivel de direcciones urbanas.

Este documento se organiza en Capítulos; en el capítulo 2 se abordan los conceptos relacionados con los Sistemas de Información Geográfica (SIG) y geocodificación a nivel de direcciones urbanas. En el capítulo 3 se describe la metodología aplicada con sus diferentes fases. En el capítulo

4 se describen los resultados obtenidos con la investigación realizada. En los capítulos 5 se discuten algunos aspectos encontrados durante el desarrollo de la investigación. En capítulo 6 se describen las conclusiones obtenidas y finalmente en el capítulo 7 se realizan algunas recomendaciones que pueden ser consideradas para futuras investigaciones.

## **TEMA DE INVESTIGACIÓN**

El tema de investigación propuesto comprende la implementación del geocodificador de direcciones urbanas para el municipio de Pasto. Este geocodificador será integrado con SIGODEP, el sistema georreferenciado del Observatorio del delito del municipio de Pasto, georreferenciando los eventos delictivos fatales y no fatales registrados por esta entidad. Esto permitirá visualizar los delitos georreferenciados a nivel de direcciones urbanas e identificar patrones visuales o zonas de riesgo donde estos ocurren con frecuencia de acuerdo a la selección y configuración previa de variables.

## **ÁREA DE INVESTIGACIÓN**

El desarrollo de la investigación involucra la apropiación y aplicación de conocimiento relacionado con las áreas de Sistemas de Información Geográfica (SIG) y gestión de Bases de Datos (DB).

## **LÍNEA DE INVESTIGACIÓN**

La presente investigación corresponde a la Línea de Software y Manejo de Información, con enfoque al procesamiento y análisis de información almacenada en bases de datos.

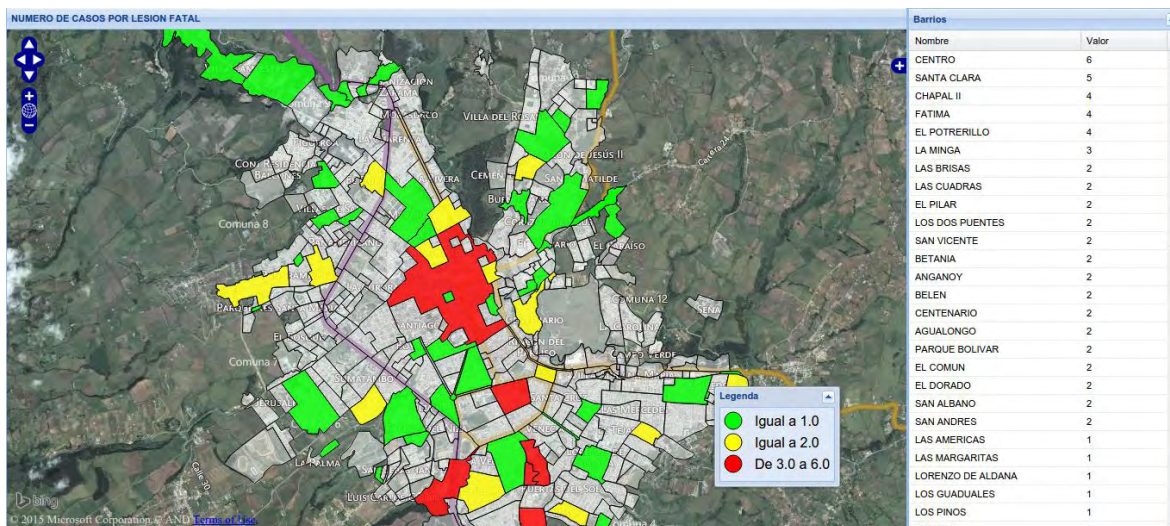
## **DESCRIPCIÓN DEL PROBLEMA**

SIGODEP es el sistema de información georreferenciado del Observatorio del Delito de Pasto bajo software libre [3], desarrollado en la primera fase del proyecto de investigación denominado "Detección de patrones delictivos con técnicas de minería de datos en el Observatorio del Delito del municipio de Pasto" financiado mediante convenio 615-2011 por COLCIENCIAS con contrapartida de la Universidad de Nariño y la Alcaldía Municipal de Pasto, y ejecutado por los investigadores del Grupo de Investigación Aplicada en Sistemas -GRIAS del departamento de Sistemas de la Universidad de Nariño.

Una de las funcionalidades que soporta este sistema es la generación de mapas temáticos, a partir de cruce de variables que describen las lesiones de causa externa fatales y no fatales que ocurrieron

en el municipio de Pasto desde el año 2002 hasta el presente año. A partir de esta funcionalidad se identifica la visualización de estadísticas correspondientes a estas incidencias desplegada en geometrías espaciales a nivel de barrios y comunas como el gránulo más pequeño de georreferenciación (véase Figura 1).

Figura 1. Georreferenciación de eventos delictivos a nivel de barrios, funcionalidad implementada en SIGEODEP.



No es posible la visualización de cada evento delictivo a nivel de direcciones urbanas ya que no se cuenta con una herramienta de libre uso que traduzca las direcciones del municipio de Pasto a coordenadas geográficas, lo cual permitiría un análisis espacial de estos eventos con mayor grado de detalle.

El desarrollo de la presente investigación permitió construir el geocodificador de código abierto para el municipio de Pasto que permita traducir las direcciones urbanas a coordenadas geográficas e integrarlo en SIGEODEP, georreferenciando las lesiones fatales y no fatales a nivel de direcciones y de esta manera permitirle al Observatorio del Delito del municipio de Pasto, visualizar zonas en el municipio donde se identifiquen situaciones atípicas o especiales sin depender de barrios o comunas como está actualmente configurado.

## FORMULACIÓN DEL PROBLEMA

Al iniciar de la investigación se plantearon los siguientes interrogantes que posteriormente fueron resueltos con el desarrollo de los objetivos del proyecto.

- ¿Cuáles son los métodos y/o herramientas apropiados para geocodificar direcciones urbanas del municipio de pasto?
- ¿De qué forma se puede obtener las direcciones urbanas del municipio de Pasto como datos de referencia?
- ¿Cuál es el proceso apropiado para geocodificar direcciones urbanas en el municipio de Pasto?
- ¿Cuáles son las técnicas y/o herramientas apropiadas para visualizar eventos delictivos a nivel de direcciones?
- ¿De qué forma se pueden verificar el funcionamiento el geocodificador de direcciones urbanas?

## **OBJETIVOS DE LA INVESTIGACIÓN**

### **General**

Implementar un geocodificador de eventos georreferenciados a nivel de direcciones urbanas que permitan identificar patrones delictivos fatales y no fatales en zonas específicas del municipio de pasto.

### **Específicos**

- Construir un repositorio de datos con las direcciones urbanas del Municipio de Pasto aplicando el estándar de la Institución Geográfica Colombiana Agustín Codazzi.
- Implementar el geocodificador de direcciones urbanas para el municipio de Pasto.
- Implementar un visor cartográfico de eventos delictivos georreferenciados a nivel de direcciones urbanas.
- Integrar el visor cartográfico con el sistema georreferenciado SIGEODEP.
- Realizar pruebas a los componentes de geocodificación y visualización de eventos delictivos y analizar los resultados obtenidos.

### **Justificación**

Una vez terminada la investigación, el Observatorio del delito del Municipio de Pasto dispone de un geocodificador de nomenclaturas urbanas que le permitirá visualizar e identificar zonas geográficas específicas donde ocurren con frecuencia delitos fatales y no fatales a través del análisis agrupado de los eventos registrados por este ente gubernamental.

Para el desarrollo de la investigación se utilizaron herramientas con licencia software libre, lo que permitió la reutilización del geocodificador construido por cualquier organización en el municipio de Pasto que requiera georreferenciar eventos registrados e identificar según su contexto funcional, zonas de interés en el municipio a partir de direcciones específicas de ocurrencia.

El geocodificador servirá como herramienta de apoyo en la investigación y determinación de decisiones frente al descubrimiento de situaciones atípicas o problemas en zonas del municipio de Pasto.

Finalmente, esta investigación permitió adquirir experiencia investigativa en el área de sistemas de georreferenciación de eventos para el Municipio de Pasto.

### **Delimitación**

El área geográfica que cubre el municipio de Pasto está conformada por 12 comunas (área urbana) y 17 corregimientos (área rural). Con el desarrollo de esta investigación se definirá el proceso de geocodificación de direcciones que estén ubicadas en el área urbana del municipio de Pasto.

Se construyó una base de datos espacial donde se recopilarán las direcciones de los predios de las comunas 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 del municipio de Pasto. Serán considerados los dos principales casos de asignación de direcciones: direcciones asignadas según la malla vial del área urbana (ej. Carrera 20 N° 13 a 45) y asignadas según la nomenclatura barrio, manzana e identificación de predio (ej. Barrio Tamasagra, Manzana 10 Casa 3). Las direcciones fueron recopiladas según la norma de estandarización de nomenclaturas urbanas propuesta por el Instituto Geográfico Agustín Codazzi de Colombia.

El geocodificador de direcciones fue integrado con el sistema georreferenciado del observatorio del delito del municipio de Pasto - SIGEODEP a través de la implementación de un visor cartográfico web, permitiendo visualizar el resultado del proceso de geocodificación de delitos ejecutado.

La georreferenciación de eventos delictivos implicó la geocodificación de Lesiones de causa externa Fatales y No fatales registradas por el observatorio del delito del municipio de Pasto, a partir de la dirección de ocurrencia registrada por cada delito. El acoplamiento del geocodificador SIGEODEP implicará la integración de la funcionalidad de conteo de eventos delictivos implementada en este sistema con la generación de mapas de puntos y de calor.

## 1. MARCO TEÓRICO

### 1.1. ANTECEDENTES

Tras el proceso de búsqueda de antecedentes que aporten al proceso de contextualización de la investigación se encontraron proyectos de investigación realizados en diferentes áreas de aplicación, en su mayoría orientadas al apoyo en la planificación territorial y a la identificación de patrones de ocurrencia de eventos relacionados con temas de orden social mediante la creación de herramientas de soporte basadas en Sistemas de Información Geográfica. A continuación se citan investigaciones destacadas:

En el ámbito internacional se destaca TIGER Geocoder, un geocodificador de códigos postales y direcciones en los Estados Unidos creado por [4]. Se basa en el Sistema Geográfico de Codificación y Referenciación Topológicamente Integrado (TIGER, por sus siglas en inglés) publicado por la Oficina del Censo de los Estados Unidos [5] para la estandarización de descripciones de objetos geográficos como calles, edificios, ríos o regiones en ese país. Fue desarrollado bajo software libre en el lenguaje procedural PL/SQL del sistema gestor de bases de datos PostgreSQL utilizando funciones de la extensión PostGIS para gestión de datos espaciales. De igual forma se destaca [4], adaptando TIGER Geocoder para la creación de Cartociudad, un sistema de geocodificación y visualización cartográfica digital que recopila información sobre la red vial, la división administrativa censal y la división postal de España. [7] identifica los procesos de Ingeniería inversa realizados para conocer, adaptar e implementar la arquitectura y componentes de TIGER Geocoder en el Sistema Gestor de Bases de Datos con soporte de datos espaciales Oracle Spatial.

En el ámbito nacional se identifica el aplicativo web Mapas.com.co desarrollado por [8], cuyo fin es brindar a usuarios u organizaciones servicios de publicación de información comercial o de negocios en la web sobre mapas digitales mediante contratos ASP (arrendamiento de servicios)[9]. Se destaca los servicios de geocodificación (a nivel de direcciones residenciales y sitios de interés) y búsqueda de rutas óptimas de recorrido con cobertura en gran parte del territorio colombiano, ambos realizados con alto grado de precisión reflejada en sus resultados. De igual forma se destaca el estudio realizado por [10], en el cual comparan los resultados arrojados por los procesos

determinístico y probabilístico de geocodificación implementados en la Alcaldía de Medellín para la georreferenciación de información primaria en todos sus procesos de gobernabilidad y desarrollo territorial. Las herramientas analizadas fueron: la funcionalidad Address Locator (denominada GEOCODING por la entidad gubernamental) provista por la suite comercial ArcGIS para gestión de información geográfica (geocodificación probabilístico) y el geocodificador desarrollado para la ciudad de Medellín denominado GEOCOD, una aplicación transversal desarrollada en el lenguaje procedural PL/SQL de Sistema Gestor de Bases de datos Oracle y el lenguaje de programación Python (geocodificación determinístico) [10].

En el ámbito regional y del departamento se identifican dos investigaciones importantes. El trabajo realizado por [11] tiene el objetivo de apoyar a la toma de decisiones de la subsecretaria de Planeación Municipal sobre el Plan de Ordenamiento Territorial del municipio de Pasto a través del desarrollo de un sistema de información geográfica web. El sistema recopila información geográfica general (comunidades, ríos, perímetros) y con detalle sobre la comuna 3 (delimitación de la comuna, sectores, predios) del municipio de Pasto. Otra investigación es la realizada por [12]. Esta herramienta fue desarrollada con el objetivo de facilitar a las pequeñas y medianas organizaciones de la región y el país administrar su información de forma georreferenciada. Permite la gestión y visualización de mapas cartográficos creados a partir de múltiples fuentes de información. Se destaca el módulo de geocodificación el cual propone un algoritmo de geocodificación de direcciones denominada Atlas Street Geocoding.

## **1.2. SUPUESTOS TEÓRICOS**

**1.2.1. Información georreferenciada.** La información es un recurso utilizado por personas y organizaciones para conocer las causas y consecuencias de la ocurrencia de eventos, fenómenos o problemas en el entorno que los rodea. Esta se compone de datos, que corresponden a la identificación aislada de las características de dichos eventos.

La información georreferenciada se define como información que contiene datos espaciales, los cuales denotan una ubicación sobre la superficie terrestre. Estos datos espaciales normalmente corresponden a pares de coordenadas que definen la posición y forma de una figura geométrica particular que representa objetos de la realidad (calles, ríos, edificios, predios, semáforos, etc.) y que pueden ser medidos, relacionados y visualizados gráficamente [13].

Este tipo de información es realmente útil en búsquedas de soluciones a problemas de orden geográfico. [14] identifica algunos de estos posibles problemas:

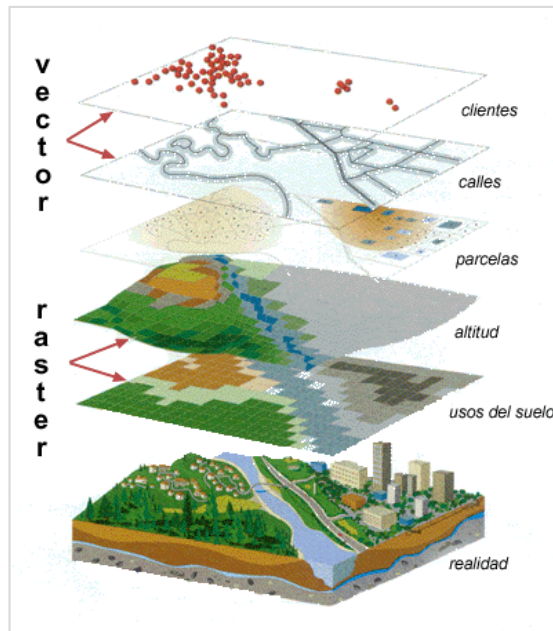
- Ubicación y georreferenciación de predios, entidades de servicios públicos, entidades gubernamentales, etc.
- Identificación de rutas para construcción de nuevas vías.
- Determinación de rutas optimas de recorrido.
- Ubicación estratégica de negocios o sucursales de ventas.
- Gestión y preservación de recursos naturales.
- Gestión y promoción de sitios turísticos y de interés (Centros comerciales, museos, parques, etc.).

En el auge del desarrollo de tecnologías de la información y la comunicación en diferentes áreas de aplicación surgen los Sistemas de Información Geográfica (SIG) y con ellos tecnologías basadas en computadoras para la captura, verificación, almacenamiento, consulta, análisis y despliegue de datos georreferenciados con ubicaciones geográficas en la tierra [15]. Los SIG proveen herramientas especializadas que permiten una gestión adecuada de información geográfica lo cual contribuye a descubrir conocimiento, soporte fundamental en la resolución de problemas y toma de decisiones.

**1.2.2. Bases de datos espaciales.** En [14] se define base de datos como la herramienta de gestión basada en software de computadoras que permite organizar y administrar la información de una organización de modo que sea posible acceder a ella y utilizarla en forma eficiente. En el caso de los SIG se denomina base de datos espacial, la cual extiende este enfoque funcional al soporte de información geográfica constituida por datos espaciales.

Las bases de datos espaciales permiten definir estructuras que soportan el almacenamiento de los componentes principales que constituyen a la información geográfica identificados por [2]: espacial (datos relacionados con la posición dentro de un sistema de referencia determinado) y temático (datos relacionados con la naturaleza de dicho fenómeno, sus características particulares). [17] afirma que este tipo de bases de datos proporcionan funciones especiales e índices para almacenar datos basándose en modelos de representación de información geográfica, gestión de información mediante capas temáticas (véase Figura 2) manipulación y consulta de datos haciendo uso de recursos como el Modelo de datos Relacional y extensiones del Lenguaje de Consulta Estructurado (SQL) para realizar consultas espaciales.

Figura 2. Gestión de información geográfica mediante capas temáticas



Fuente: [www.geogra.uah.es/gisweb/1modulosespanyol/IntroduccionSIG](http://www.geogra.uah.es/gisweb/1modulosespanyol/IntroduccionSIG)

**1.2.3. Modelos de representación geográfica.** La gestión de datos georreferenciados considera métodos o modelos para representar los objetos geométricos presentes en el mundo real (calles, vías, edificios, puentes, etc.). Existen dos modelos de representación o formatos fundamentales: Vectorial y Raster. La Figura 4 muestra la forma como estos modelos representan objetos geográficos.

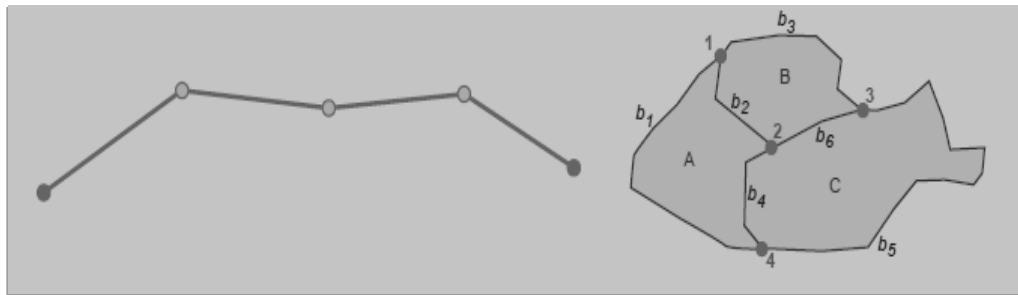
**Modelo Vectorial:** En este modelo los objetos del mundo real se representan mediante puntos, líneas y polígonos que definen sus límites o fronteras a partir de un sistema de coordenadas definido previamente para localizar cada objeto. [1] identifican los objetos fundamentales utilizados en este modelo. En la Figura 3 se muestra la representación geométrica de estos elementos.

- Puntos: Son objetos que se definen con pares de coordenadas (x, y) en planos 2D o coordenadas triples (x, y, z) en planos 3D. Se utilizan para ubicar características de una sola dimensión (direcciones urbanas, sitios de interés, sitios turísticos entre otros).
- Líneas: Son usadas para representar objetos de una sola dimensión como caminos, ferrocarriles, canales, ríos, líneas de alta tensión. La mayoría de estos objetos son

representados mediante líneas curvas, las cuales se estructuran mediante nodos. Cada par de nodos forma líneas contiguas, las cuales se denominan segmentos.

- Polígonos (Áreas): Son utilizados para representar aproximaciones de objetos o áreas geográficas con límites curvilíneos. Puede tratarse de líneas curvas en donde se juntan sus extremos o áreas más complejas como áreas divididas en sub-áreas.

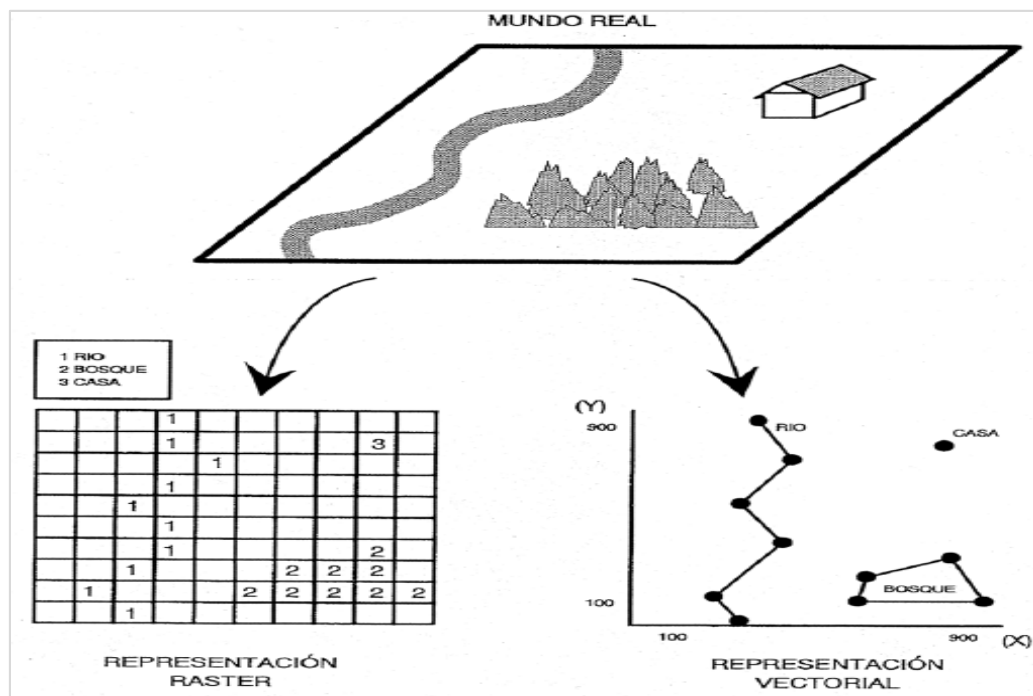
Figura 3. Representación objetos geográficos: línea y polígonos (áreas).



Fuente: Principles of Geographic Information Systems: An Introductory Textbook Hussman. 2009

**Modelo Raster:** De acuerdo con [2] En este modelo, el espacio está representado por un conjunto de celdas adyacentes llamadas píxeles, que representan las unidades de información espacial. Estas establecen su localización por un sistema de referencia en filas y columnas, acompañado por la extensión del mapa y el tamaño de la celda. Los píxeles en realidad no mantienen una relación mutua entre sí. En la cobertura de tipo Raster, cada celda tiene un valor o código asignado, correspondiente al tipo de información temática que representa la celda.

Figura 4. Representación de la realidad en los formatos Vectorial y Raster.

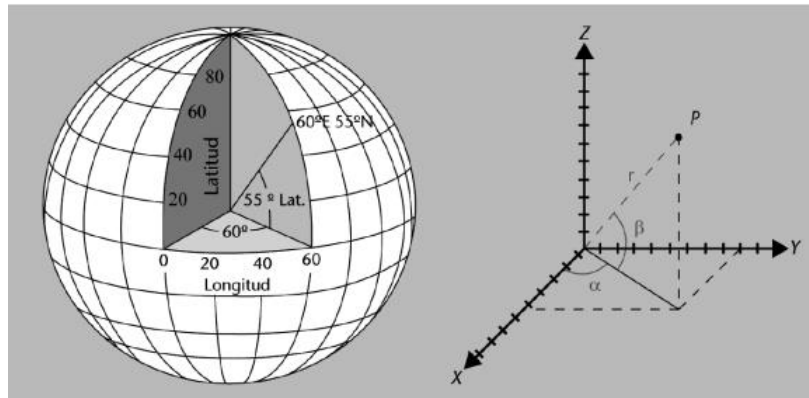


Fuente: Sistemas de Información Geográfica Versión 1.0. Olaya, 2011.

**1.2.4. Sistemas de referencia espacial.** La ubicación de objetos en la superficie terrestre obedece a sistema de referencia espacial los cuales mediante referencias geométricas permiten medir su ubicación. Se identifican dos principales sistemas: El sistema de referencia geodésico y el sistema de proyección cartográfica.

**Sistema de referencia geodésico:** Este sistema considera la superficie terrestre como una figura esférica sobre el cual se proyectan unidades de medida angulares denominadas coordenadas geográficas. Estas coordenadas se obtienen mediante operaciones matemáticas utilizando dos ángulos medidos desde el centro de la Tierra, denominados latitud y longitud. La longitud corresponde al ángulo medido a lo largo del ecuador entre el meridiano de Greenwich (longitud 0°) y cualquier punto de la tierra, la latitud corresponde al ángulo medido entre la línea Ecuatorial y la posición de cualquier punto sobre la superficie terrestre (véase Figura 5).

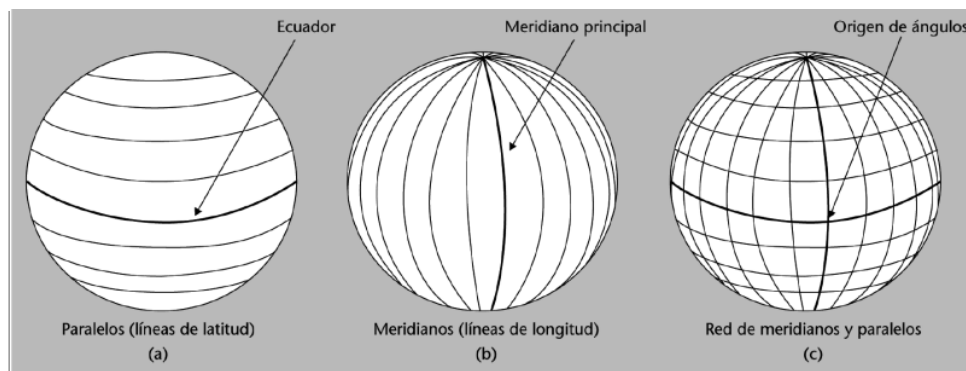
Figura 5. Valores de longitud y latitud de un punto sobre la esfera.



Fuente: Introducción a los sistemas de información geográfica y geo telemática. Navarro, 2011.

Las líneas formadas a partir de las longitudes son denominadas meridianos, el meridiano principal denominado Ecuador o meridiano de Greenwich divide la esfera terrestre en dos hemisferios, oriental y occidental. Las líneas formadas por las latitudes se denominan paralelos, el paralelo principal o Ecuador divide la esfera terrestre en dos hemisferios, norte y sur (véase Figura 6).

Figura 6. Elementos de los sistemas de coordenadas geográficas: Paralelos (a), Meridianos (b) Origen de coordenadas (c)

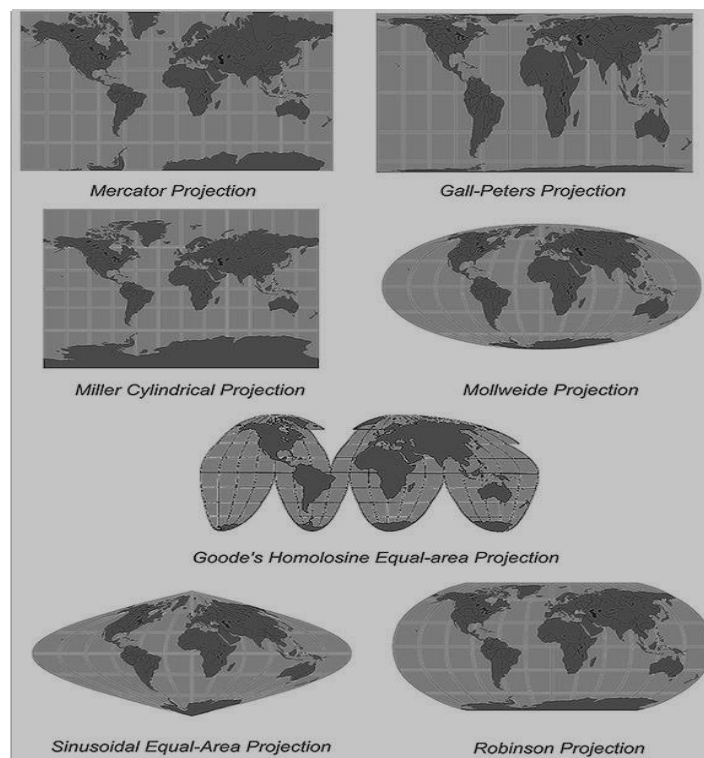


Fuente: Introducción a los sistemas de información geográfica y geo telemática. Navarro, 2011.

**Sistema de proyección cartográfica:** Dado que los meridianos y los paralelos son ubicados según el sistema de referencia geodésico sobre la superficie esférica terrestre, deben considerarse métodos matemáticos adicionales que permitan proyectar estos elementos sobre una superficie plana manteniendo su exactitud posicional. Según [18], las dimensiones de las coordenadas, las distancias, los ángulos entre paralelos y meridianos, las áreas que estas encierran y las direcciones, son posibles de traspasar al plano, pero nunca conjuntamente en una misma proyección.

En la búsqueda de una mejor representación de la realidad para determinadas áreas o para toda la superficie de la Tierra existen sistemas de construcción de proyecciones sobre superficies planas que tratan de representar al menos una de estas cualidades de la realidad en forma correcta y sin distorsión. En la Figura 7 se muestra algunas de estas proyecciones.

Figura 7. La superficie terrestre representada según distintas proyecciones

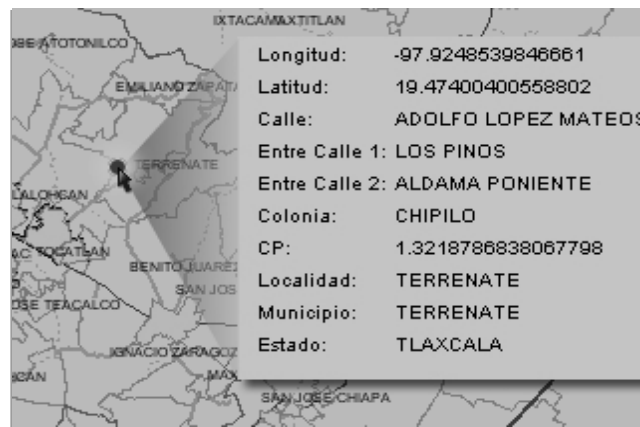


Fuente: [geology.isu.edu/geostac/Field\\_Exercise/topomaps/distortion.htm](http://geology.isu.edu/geostac/Field_Exercise/topomaps/distortion.htm)

Dentro de la gestión de información geográfica es indispensable establecer un sistema de proyección adecuado que se aplique al despliegue y visualización correcta de objetos geográficos y sus respectivas coordenadas geográficas.

**1.2.5. Geocodificación de eventos.** La geocodificación como proceso permite asignar coordenadas geográficas a los datos temáticos de eventos mediante el procesamiento de objetos espaciales asociados a esta información. Estas coordenadas pueden ser utilizadas para desplegar, localizar y visualizar con detalle estos eventos en mapas cartográficos digitales. En la Figura 8 se muestra un ejemplo de eventos geocodificados desplegando sus respectivas características.

Figura 8. Geocodificación de eventos y despliegue de datos temáticos o características.



Fuente: [www.mapdata.com.mx/servicios\\_geocodificacion.html](http://www.mapdata.com.mx/servicios_geocodificacion.html)

Realizar este proceso con alto grado de aproximación a la realidad geográfica se considera como un objetivo fundamental que conlleva a considerar aspectos sobre este proceso. La información cartográfica debe tener suficiente detalle sobre el área geográfica de estudio que permita identificar con precisión la ubicación que se está buscando.

Para cumplir con este objetivo, [19] identifica tres componentes básicos que requiere el proceso de geocodificación:

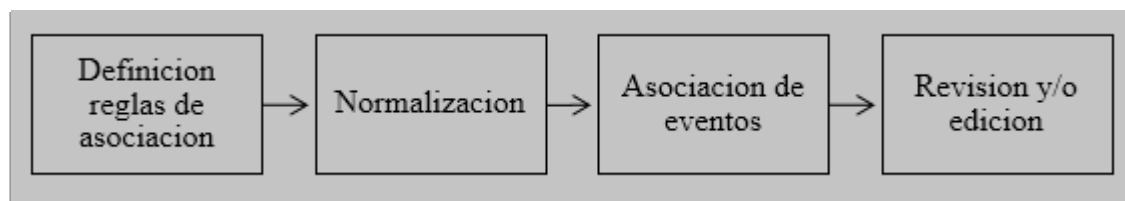
- **Datos Descriptivos:** Se refiere a registros que describen eventos ocurridos en un entorno geográfico y que, por lo tanto, contienen alguna característica que permite georreferenciarlos en dicho entorno. Son capturados por entes organizacionales generalmente sin tener en cuenta alguna norma o formato lo que conlleva a obtener registros con errores y omisiones. Sin embargo, estos errores pueden ser corregidos mediante un proceso de normalización o búsqueda en bases de datos de referencia.
- **Datos de Referencia:** Son datos que describen eventos de referencia presentes sobre un entorno geográfico y que son recopilados desde diferentes fuentes de información según

estructuras o estándares de georreferenciación establecidas. A diferencia de los datos descriptivos, los eventos descritos mediante datos de referencia son relacionados a elementos que pueden ser representados en el espacio, desde coordenadas geográficas hasta elementos complejos como estructuras de información espacial.

- **Software:** Es la herramienta basada en hardware y software especializado que asocia datos descriptivos con datos de referencia obteniendo como resultado la codificación espacial de eventos y sus características. La implementación de esta herramienta debe estar orientada a uno o varios criterios de geocodificación según se estipulen los requerimientos de información de una organización (geocodificación por barrios, manzanas, direcciones, sitios de interés, entre otros).

**Proceso de geocodificación:** Una vez estructurados los elementos necesarios para realizar el proceso de geocodificación se debe definir la forma de interacción de estos elementos dentro de dicho proceso. En la Figura 9 se muestra el proceso general de geocodificación de eventos expuesto por [19].

Figura 9. Proceso general de geocodificación



Fuente: ArcGIS 9, Geocoding Rule Base Developer Guide. ESRI, 2003.

Las reglas de asociación de eventos definen aspectos del método de geocodificación a utilizar como la estructura de la base de datos, atributos sobre los cuales se realizara la geocodificación, atributos que deben indexarse para optimización de búsquedas, manejo de errores, criterios de evaluación de resultados, entre otros. La normalización hace referencia a describir los atributos georreferenciados de los datos descriptivos y los datos de referencia ajustándose a normas o formatos de abreviaturas o escritura. En la asociación de eventos el software compara cada registro descriptivo con cada registro de referencia mediante el atributo seleccionado generando puntajes que determinan el grado de similitud de estos dentro de un rango de aceptación establecido. De acuerdo a estas revisiones se ajustan las reglas definidas para obtener resultados acordes a los requerimientos de información establecidos según el caso de estudio.

**Tipos de asociación de información:** Uno de los objetivos evidentes de la asociación de eventos es integrar diferentes fuentes de información que corresponden a entidades únicas como personas, objetos, lugares, etc. En muchos casos estas fuentes poseen formatos o estructuras diferentes lo cual dificulta relacionarlas. Se deben considerar tipos de asociación o enlace que pueden realizarse los registros de estas fuentes.

- **Método determinístico o exacto (Deterministic Record Linkage):** Es la aplicación de técnicas o procedimientos orientados a determinar las coincidencias exactas entre registros de los datos descriptivos y los datos de referencia a partir de un atributo en común. Según lo afirmado por [20], un ejemplo claro del método determinístico o exacto en bases de datos relacionales es la vinculación de registros por medio de claves primarias y foráneas sin importar su tipo, ya sean numérico o compuesto por caracteres. En el caso de geocodificación de direcciones urbanas se comparan los datos que corresponden al atributo dirección dentro de los datos descriptivos y en los datos de referencia obteniendo relaciones o vínculos exactos.
- **Método Probabilístico (Probabilistic Record Linkage):** Es la aplicación de técnicas o procedimientos orientados a encontrar similitud entre dos datos bajo condiciones de incertidumbre. Este método es utilizado en casos donde existan fuentes de información con atributos que hacen referencia a una misma característica pero que se describen con diferentes formatos de escritura o simplemente no se rigen de alguno. Como lo afirma [21], La relación de registros debe ajustarse al manejo de errores de comparación determinados por la probabilidad de coincidencia del registro calculada a partir de algoritmos matemáticos. La asignación y evaluación de tales probabilidades tiene la intención de imitar al ser humano en la determinación de registros similares.

**1.2.6. Visualización de información geográfica.** Visualizar la información geográfica según afirma [2], es una parte fundamental donde se identifican muchas de las ventajas de trabajar con SIG. Es posible encontrar SIG enfocados al análisis en los cuales no existe forma de visualizar la información con la que se trabaja, sin embargo en su mayoría estos incluyen funcionalidades de visualización como elemento básico de representación de información geográfica.

[22] identifica las ventajas que se tiene al manejar información geográfica haciendo uso de herramientas cartográficas basadas en SIG

- Ahorro de tiempo en procesos manuales.
- Manejo de grandes volúmenes de datos y almacenamiento en formatos compactos.

- Flexibilidad en las opciones de salida. Generación de diferentes productos a partir de un solo sistema.
- Generación y manejo de bases de datos espaciales basados en modelo de datos.
- La información geográfica (tanto temática como espacial) puede ser extraída y manipulada en forma simultánea e interrelacionada.
- Funciones como recolección de datos, análisis espacial y toma de decisiones son integradas dentro de un contexto común.
- La automatización en el manejo de información geográfica permite mantener la información geográfica, y por ende, los mapas actualizados.

**Cartografía web:** La cartografía web se considera una tendencia importante que ha tomado la cartografía tradicional. Hoy en día se cuenta con servicios web de bajo costo y de libre acceso a información cartográfica lo cual permite tanto a personas como organizaciones interesadas crear y mejorar mapas de la web de acuerdo a necesidades particulares de información.

[23] propone una clasificación de mapas web de la cual se destacan los siguientes tipos:

- a) **Cartografía web estática:** Cartografía publicada sin opciones que permitan su manipulación (ej. Mapas escaneados). A pesar de su desventaja con respecto a interactividad, en muchos casos posee información de referencia importante para la creación de cartografía interactiva.
- b) **Cartografía web híbrida o distribuida:** Aquella que es generada a partir del acceso a una o varias fuentes de información provenientes de servicios distribuidos. Esto podría incluir la extracción de datos de otros sitios web que no fueron empaquetados expresamente con fines cartográficos.
- c) **Cartografía web interactiva:** Cartografía publicada con opciones de configuración tales como:
  - Alternar la activación y desactivación de las capas del mapa.
  - Permitir que el usuario pueda configurar variables de despliegue según un determinado contexto de análisis.
  - Obtener información detallada sobre las características del mapa.
  - Producir cartografía personalizada.

- Explorar patrones de datos y relaciones usando herramientas de visualización que están dinámicamente relacionados con el mapa.

La característica interactiva se logra en gran parte a la interacción directa de usuarios con las estructuras de almacenamiento de datos y con los datos espaciales utilizados para el despliegue de los mapas cartográficos.

- d) **Cartografía web colaborativa:** En muchos países, a excepción de los Estados Unidos, el acceso a información geográfica es restringido para cualquier usuario, mapas que en apariencia son gratuitos manejan restricciones técnicas o legales para su uso, prohibiendo a las personas su utilización de manera creativa o productiva. En contraste a esta situación surge la cartografía web colaborativa, la cual permite a cualquier usuario la libre colaboración en la creación y modificación de mapas cartográficos y el libre acceso a esta información.

Uno de los proyectos más conocidos es Open Street Maps. [24] lo define como un proyecto colaborativo cuyo objetivo es crear un mapa mundial editable al cual puede acceder cualquier ciudadano interesado en agregar, modificar o extraer datos espaciales sobre algún área geográfica de interés.

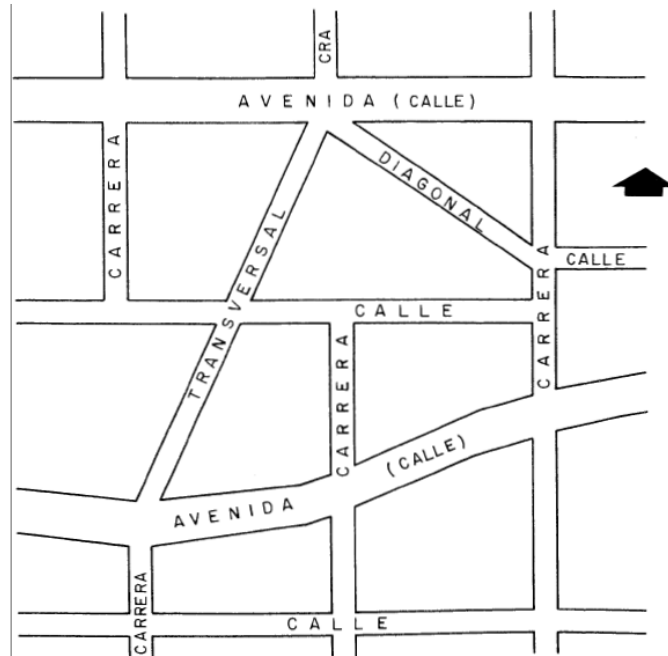
**1.2.7. Direcciones urbanas.** Según [25], la nomenclatura urbana (popularmente denominada dirección) es la identificación tanto de vías como de predios que conforman el área de una ciudad o población, con signos numéricos y alfanuméricos. La asignación de nomenclaturas urbanas permite a personas y organizaciones localizar lotes de terreno o viviendas mediante un sistema de planos y letreros que indican los números o los nombres de las calles, manzanas catastrales y/o edificios. Las nomenclaturas son indispensables en el ordenamiento urbano de una ciudad.

[26] identifica los elementos presentes en áreas urbanas que requieren o no asignación de nomenclaturas. En la figura 10 se muestra la distribución de cada elemento.

- Calle: Vía pública con orientación predominante definida de acuerdo al modelo de una ciudad.
- Carrera: Vía pública generalmente perpendicular a la calle. Genera nomenclatura predial.
- Diagonal: Vía pública que generalmente tiene el mismo sentido que la calle sin ser paralela a esta. Puede o no genera nomenclatura predial.
- Transversal: Vía pública que generalmente tiene el mismo sentido que la calle sin ser paralela a esta. Puede o no generar nomenclatura predial.

- Avenida: Vía pública cuyas especificaciones y características son notoriamente superiores a las vías predominantes.

Figura 10. Elementos del área urbana que requieren asignación de nomenclaturas



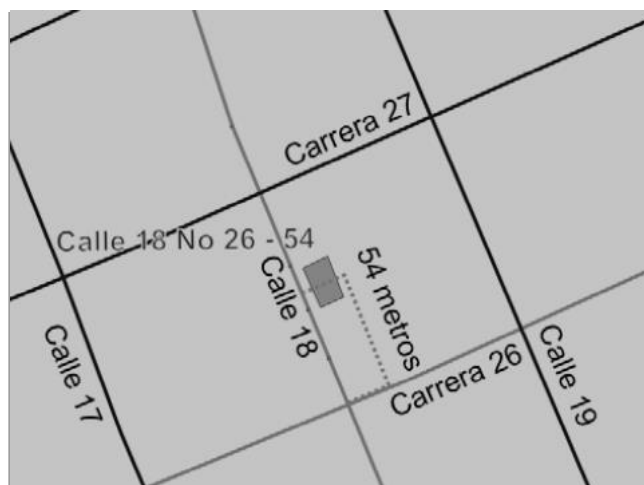
Fuente: Manual de reconocimiento predial. Instituto Geográfico Agustín Codazzi, 1997.

**Direcciones urbanas en el municipio de Pasto:** Para el municipio de Pasto el ente encargado de la gestión, asignación y mantenimiento de la nomenclatura urbana es el Instituto Geográfico Agustín Codazzi – IGAC [26]. Las calles de la ciudad son enumeradas en sentido occidente a oriente y las carreras las vías en sentido sur a norte. Según su tipo una vía se nombra con la palabra calle o carrera seguida de un número arábigo y en ciertos casos de una letra, por ejemplo, cuando se abre una vía en medio de un par de vías existentes. La numeración crece indefinidamente en sentido positivo, sin embargo, de requerir nombrar vías con un número inferior a 1, al no emplearse el número cero ni números negativos las vías se nombran aumentando las palabras este o sur según corresponda el sector de la ciudad.

Una cuadra se origina de la intersección entre una vía y otras dos vías de orientación opuesta, por ejemplo, la cuadra formada en la calle 18 entre carreras 26 y 27, esta cuadra se denomina calle 18 N° 26. Para nombrar una propiedad se coloca el nombre de la cuadra en la que se encuentra, para

diferenciar los predios dentro de la misma cuadra se agrega un número correspondiente a la distancia en metros desde la puerta de la propiedad hasta la esquina formada por la intersección con la vía de inferior valor. La Figura 11 muestra un ejemplo de asignación de nomenclatura. Normalmente el último número es para para indicar casas a la derecha de la vía e impar para indicar casas a la izquierda. En este caso para determinar el sentido izquierdo o derecho debe recorrerse la vía en el sentido de avance de las vías de orientación contraria.

Figura 11. Asignación de direcciones en el área urbana del municipio de pasto



Fuente: Manual de reconocimiento predial. Instituto Geográfico Agustín Codazzi, 1997.

**Propuesta de estandarización de direcciones urbanas:** [25] propuso un estándar de direcciones urbanas que incluye la mayor cantidad de casos a nivel nacional, presentando 5 casos de asignación de direcciones que permiten un almacenamiento organizado de esta información. Dicha propuesta se basa en la organización de campos del Código Unificado de Nomenclatura Urbana (CUNU, por sus siglas) propuesto por el Departamento Administrativo de Catastro Distrital, en la estandarización de abreviaturas de la Circular 300 de 2001 del Instituto Geográfico Agustín Codazzi, de la Resolución 166 del 2004 del Ministerio de Educación Nacional y en el proceso de geocodificación expuesto por [19].

**Estandarización de abreviaturas:** El estándar de abreviaturas se basa en la Circular 300 de 2001 del Instituto Geográfico Agustín Codazzi y en la Resolución 166 del 2004 del Ministerio de Educación Nacional, debido a que la estructura de campos definidos en el código CUNU delimita únicamente 2 caracteres para el uso de abreviaturas. En la tabla 1 se muestran las abreviaturas

estandarizadas correspondientes a las utilizadas en la asignación de direcciones urbanas en el municipio de Pasto:

Tabla 1. Estándar de las principales abreviaturas utilizadas en el municipio de Pasto.

Nomenclatura	Abreviatura	Nomenclatura	Abreviatura
Avenida	AV	Etapa	ET
Avenida Calle	AC	Este	ESTE
Barrio	BR	Lote	LT
Bis	BIS	Manzana	MZ
Bloque	BQ	Norte	NORTE
Calle	CL	Oeste	OESTE
Carrera	KR	Parque	PQ
Casa	CS	Pasaje	PJ
Centro Comercial	CE	Sur	SUR
Diagonal	DG	Transversal	TV
Edificio	ED	Urbanización	UR
Esquina	EQ		

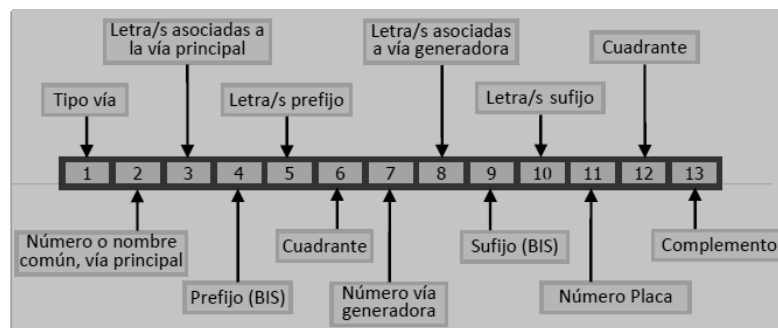
Fuente: Propuesta de Estandarización de direcciones urbanas. Instituto Geográfico Agustín Codazzi, 2009.

Según [25], la organización hace referencia al orden y la forma como deben ser capturados y almacenados cada uno de los elementos que componen una dirección urbana. Dado que no necesariamente un registro de dirección debe contener cada uno de los elementos posibles, por ejemplo no todas las vías poseen el prefijo BIS o no todas las vías deben tener explícitamente el cuadrante (Norte, Sur, Este, Oeste), estos elementos deben ser obviados dejando un espacio con el elemento siguiente, sólo y únicamente en el caso que la dirección no contemple esta información.

Para el desarrollo del geocodificador planteado se presenta los casos de asignación de direcciones de mayor aplicación en la zona urbana del municipio de Pasto.

- **Caso 1. Nomenclatura basada en la malla vial:** Sistema de nomenclatura urbano conformado básicamente por una vía principal, una vía generadora y un número de placa. Es el sistema más generalizado dentro de los centros poblados del país. En la Figura 12 se muestra la organización de estos elementos. La descripción de los mismos se muestra en la tabla 2.

Figura 12. Organización de elementos de direcciones basadas en la malla vial.



Fuente: Propuesta de Estandarización de direcciones urbanas. Instituto Geográfico Agustín Codazzi, 2009.

Tabla 2. Descripción de elementos que componen una dirección basada en la malla vial.

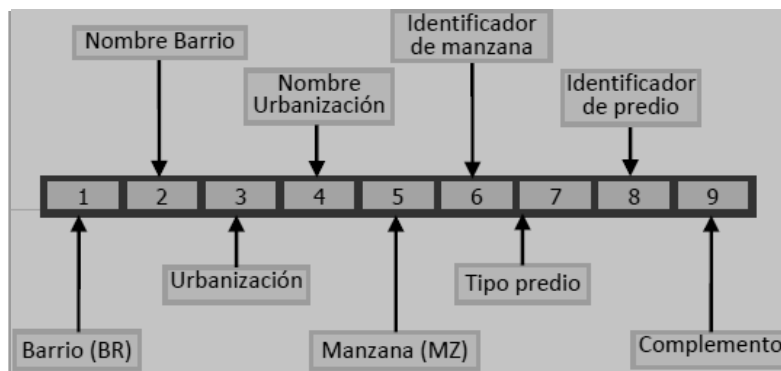
ELEMENTOS O COMPONENTES DE LA DIRECCION		
POSICION	TAMAÑO	DESCRIPCION
1	3	Tipo de vía
2	-	Numero o nombre común de la vía principal
3	-	Letra que acompaña a la vía principal
4	3	Prefijo (BIS)
5	-	Letra que acompaña al prefijo BIS
6	5	Cuadrante (NORTE, SUR, ESTE, OESTE)
7	3	Número de la vía generadora
8	-	Letra que acompaña a la vía generadora
9	3	Sufijo (BIS)
10	-	Letra, letra-letra o letra-número que acompaña al sufijo BIS
11	3	Número de la placa
12	5	Cuadrante (NORTE, SUR, ESTE, OESTE)
13	-	Complemento (PISO, LOCAL, OFICINA, MANZANA, etc.)

Fuente: Propuesta de Estandarización de direcciones urbanas. Instituto Geográfico Agustín Codazzi, 2009.

- **Caso 2. Nomenclatura basada en la estructura Barrio – Manzana – Predio:** Sistema de nomenclatura urbano que puede ser complemento del sistema de nomenclatura presentado en

el caso 1. También puede comportarse como un sistema independiente, debido generalmente a la no designación de nomenclatura a la malla vial circundante. Está conformado por: Un nombre de barrio, un código de manzana y un código de predio, aunque en la práctica se presentan más casos, que surgen como consecuencia de la combinación de distintos elementos. En la figura 14 se muestra la organización de estos elementos. La descripción de los mismos se muestra en la tabla 3.

Figura 13. Organización de elementos de direcciones basadas en la estructura Barrio – Manzana - Predio.



Fuente: Propuesta de Estandarización de direcciones urbanas. Instituto Geográfico Agustín Codazzi, 2009.

Tabla 3. Descripción de elementos de direcciones basadas en la estructura Barrio, Manzana Identificación de Predio.

ELEMENTOS O COMPONENTES DE LA DIRECCION		
POSICION	TAMAÑO	DESCRIPCION
1	2	Barrio (BR)
2	-	Nombre del barrio
3	2	Urbanización (UR), SuperManzana (SM), bloque (BQ), torre (TO), célula (CE), sector (SC), etapa (ET), ciudadela (CD).
4	-	Nombre o identificación del elemento anterior.
5	2	Manzana (MZ), interior (IN), etc.
6	-	Nombre o identificación del elemento anterior.
7	2	Tipo de unidad predial: casa (CS), lote (LO), apartamento (AP), etc.
8	-	Identificación del predio.
9	-	Complemento (Vía principal, distancia, etc.)

Fuente: Propuesta de Estandarización de direcciones urbanas. Instituto Geográfico Agustín Codazzi, 2009.

**1.2.8. Caso de estudio Observatorio del delito del municipio de Pasto.** [27] es una dependencia de la secretaria de gobierno que apoyada por la Secretaria de Salud y Transito del municipio, monitorea el comportamiento ciudadano a través de cuatro líneas de vigilancia: Lesiones de Causa Externa fatales, lesiones de causa externa no fatales, violencia interpersonal en familia, encuestas de victimización y percepción de seguridad mediante el análisis de los eventos delictivos registrados por diversas fuentes de información, sirviendo de apoyo a la creación de políticas para prevenir la ocurrencia de delitos. La tabla 4 muestra los eventos delictivos asociados a las lesiones de causa externa Fatales y No fatales que serán procesadas por el geocodificador de direcciones urbanas.

Tabla 4. Tipos de eventos delictivos asociados a las Líneas de Vigilancia del Observatorio del delito del municipio de Pasto

<b>Lesiones fatales</b>	<b>Lesiones no fatales</b>
Homicidios	Violencia Interpersonal en comunidad
Suicidios	Violencia Auto Infringida
Muertes por accidente de transito	Accidentes de transito
Lesiones no intencionales	Violencia no Intencional
	Violencia Interpersonal en Familia

Fuente: SIGEODEP: Un primer paso para la Detección de Patrones Delictivos con Técnicas de Minería de Datos (2012).

## **2. METODOLOGÍA**

### **2.1. TIPO DE INVESTIGACIÓN**

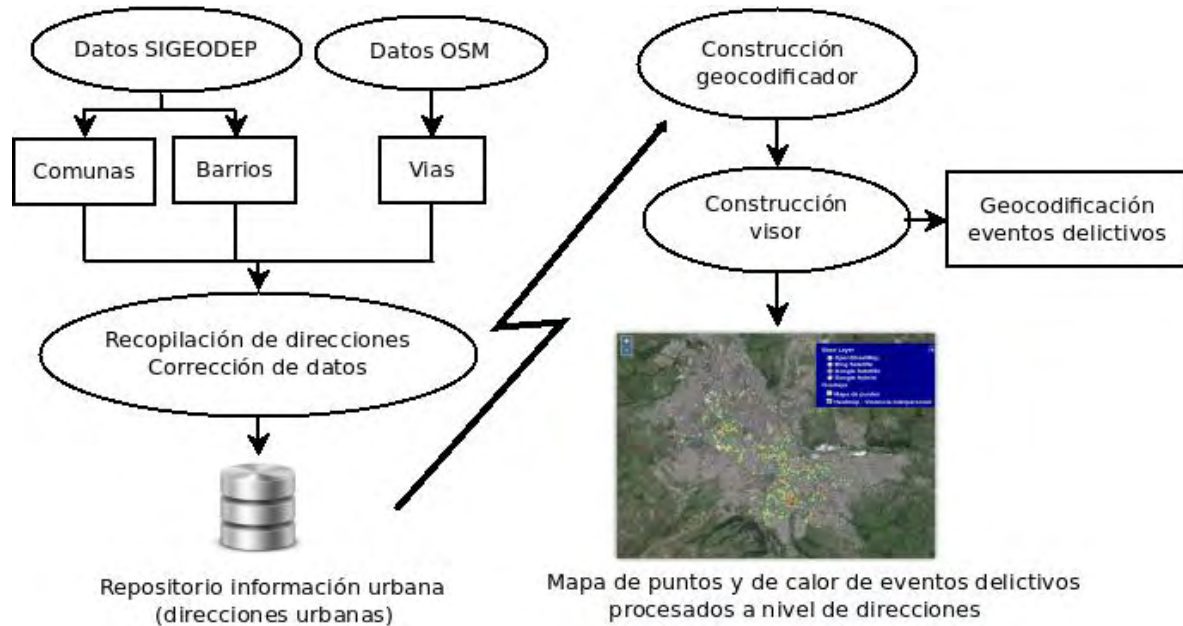
Para el cumplimiento de los objetivos propuestos se planteó una investigación de tipo **exploratoria** dado que el tema propuesto ha sido poco explorado en el contexto de estudio del municipio de Pasto. Este tipo de investigación permite seguir una metodología flexible a las diferentes pruebas de técnicas y herramientas que puedan utilizarse en el desarrollo de la investigación. Se pretende formalizar una base teórica y herramientas que permitan la formulación de futuros temas de investigación orientados a la georreferenciación de eventos a nivel de direcciones urbanas.

### **2.2. DISEÑO DE LA INVESTIGACIÓN**

Para el desarrollo de la investigación inicialmente se realizó la apropiación de conocimiento correspondiente con la temática y se definió el marco conceptual sobre el cual se iba a realizar el proyecto; como primer paso fue importante entender la problemática que se pretende solucionar, identificar los recursos necesarios para el estudio y fijar las metas que se pretende alcanzar. Para el cumplimiento de los objetivos del proyecto se planteó las siguientes fases (Véase Figura 14):

- A. Construcción del repositorio de información urbana
- B. Construcción del geocodificador de direcciones urbanas.
- C. Construcción del visor cartográfico.
- D. Definición y ejecución de pruebas.

Figura 14. Metodología aplicada en la investigación

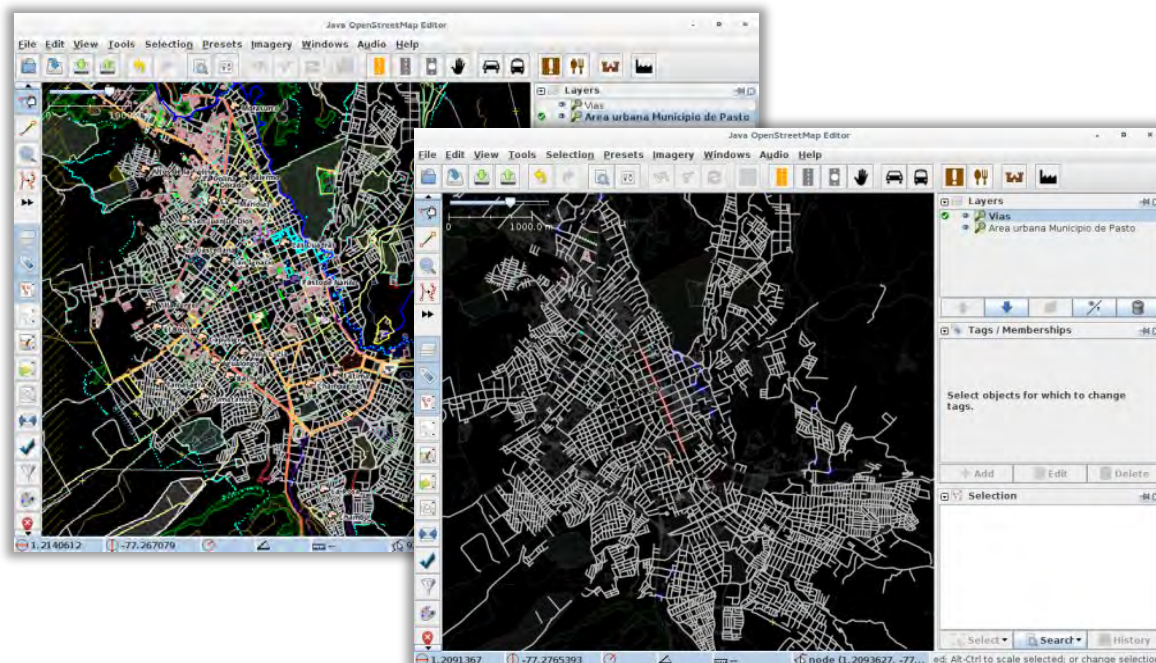


## 2.3. DESARROLLO DE LA INVESTIGACIÓN

**2.3.1. Construcción del repositorio de información urbana.** Esta fue una fase fundamental en el desarrollo de la investigación, donde se recopiló la información urbana del municipio de Pasto, incluidas las direcciones urbanas que serán utilizadas posteriormente por el geocodificador de direcciones urbanas. La creación del repositorio de direcciones urbanas digitalizadas en archivos OSM a una base de datos creada en PostgreSQL se explica con detalle en el Anexo B.

**Obtención de datos espaciales:** Inicialmente se obtuvo la capa de datos espaciales vectoriales que corresponden a la malla vial del área urbana del Municipio de Pasto, disponibles en el servicio de edición y compartición libre de mapas cartográficos en la web Open Street Maps [28]. Para la descarga de estos datos fue utilizada la herramienta de edición de datos cartográficos Java Open Street Maps Editor [29]. Se obtuvo un archivo con extensión OSM que almacena información georreferenciada del área urbana del municipio de Pasto, bajo una estructura de datos XML (Extended Markup Language). A partir de estos datos fueron filtrados los datos que corresponden a la malla vial del municipio (véase Figura 15).

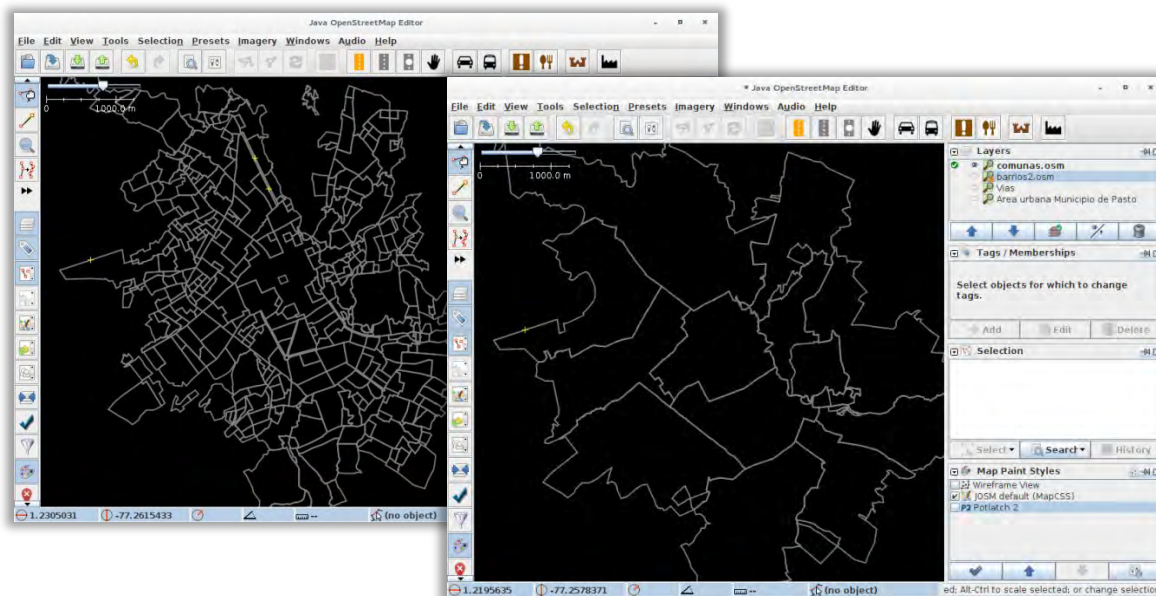
Figura 15. Lado derecho: Mapas de referencia descargados del servidor Open Street Maps. Lado izquierdo: Malla vial del municipio de Pasto filtrada de los datos descargados.



Las capas de datos espaciales que corresponden a los barrios y comunas se obtuvieron del sistema georreferenciado SIGODEP. Estas capas servirán posteriormente para asignar a cada dirección su correspondiente barrio y comuna. (Véase Figura 16).

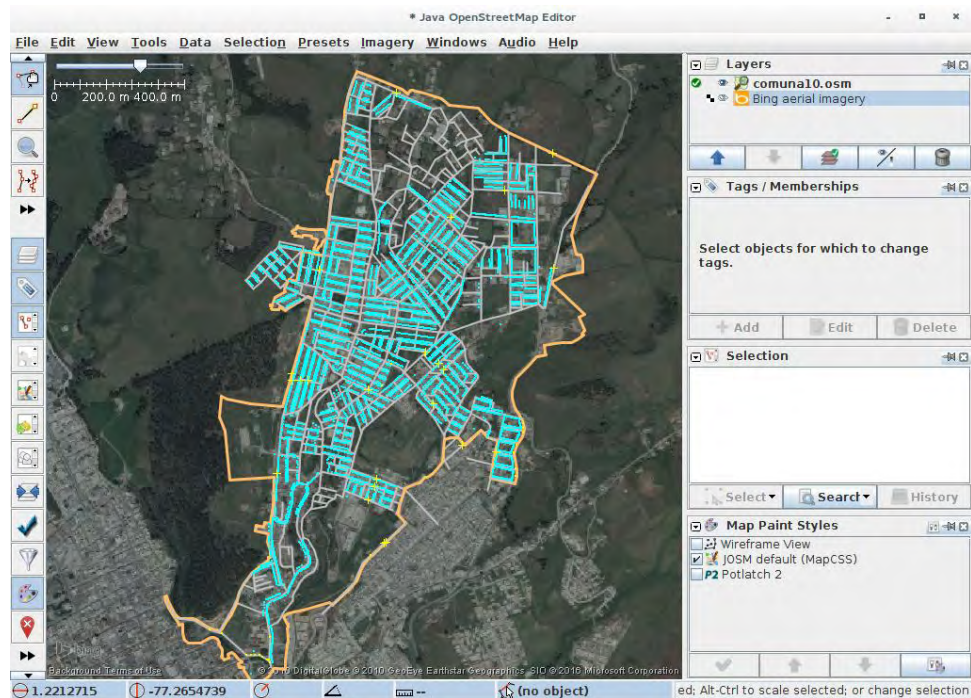
Los datos espaciales contenidos en cada archivo se almacenan en objetos geométricos basados en líneas. Cada objeto geométrico posee sus correspondientes atributos de identificación, como también las coordenadas geográficas que lo conforman. Las coordenadas asociadas a cada objeto geométrico están configuradas según el sistema de referencia espacial WGS84 (EPSG: 4326) [30].

Figura 16. Capas que corresponden a los barrios (lado izquierdo) y las comunas (lado derecho) del municipio de Pasto



**Corrección de datos obtenidos y recopilación de direcciones urbanas:** Dado que gran parte de la información geográfica obtenida se encontraba incompleta, fue realizada una previa corrección de estos datos. Al mismo tiempo, se dio inicio a la recopilación de las direcciones urbanas del municipio de Pasto. Estas tareas se ejecutaron en conjunto a medida que se recorría una a una las cuadras del municipio de Pasto. Para una recopilación ordenada de los datos se definió segmentar la malla vial por cada comuna del municipio de Pasto (véase Figura 17).

Figura 17. Segmentación de las comunas para la recopilación de direcciones. Ejemplo: Comuna 10.



Como fuente de información de referencia para corregir estos datos se utilizó el visor de mapas Google Street View [31], el visor de mapas cartográficos Bing Maps, el geoportal de visualización de datos geográficos y catastrales del Instituto Geográfico Agustín Codazzi - IGAC [30] y herramientas libres de edición de mapas cartográficos como Java Open Street Maps Editor [29] y Quantum GIS [33]. Las direcciones digitalizadas corresponden a las cuadras (o zonas) que las fuentes de información seleccionadas permitían observar.

Para la recopilación las direcciones urbanas fue utilizado el complemento Address Interpolation [34] del editor Java Open Street Maps Editor (véase Figura 18), el cual permitió generar puntos por cada cuadra recorrida de forma rápida, configurando los atributos necesarios de acuerdo con las normas de asignación de direcciones urbanas manejados en el municipio de Pasto. Se tuvo como referencia la propuesta de estandarización de direcciones del Instituto Geográfico Agustín Codazzi – IGAC [25], lo que permitió obtener direcciones con las nomenclaturas estándar que serán procesadas por el geocodificador.

Por cada punto creado, se recopilan de forma implícita atributos de georreferenciación: su correspondiente longitud y latitud. La obtención de los puntos se almacenó dentro de los archivos

OSM obtenidos inicialmente junto con los datos correspondientes a vías barrios y comunas corregidos.

Figura 18. Address Interpolation: Complemento de la herramienta Java Open Street Maps utilizado para la recopilación de direcciones urbanas.



**Definición del modelo de datos:** Luego de obtener la información urbana del municipio de Pasto corregida, se definió el modelo de datos para el almacenamiento de esta información. Se definió el modelo Entidad – Relación de la base de datos, considerando los atributos disponibles por cada tipo de datos recopilados:

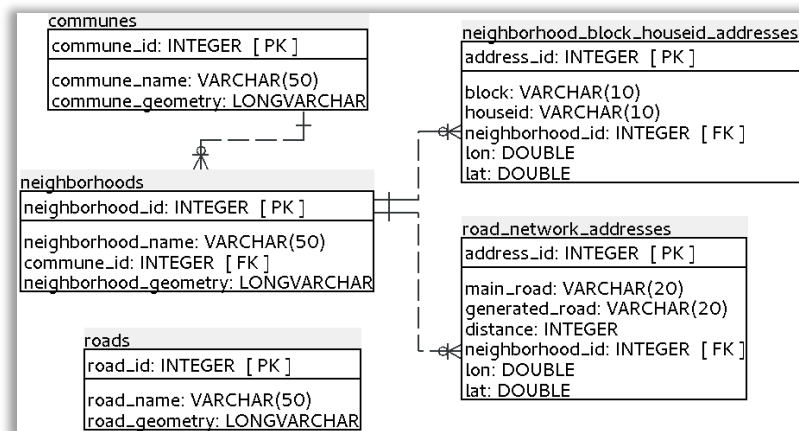
- Comunas (communes)
- Barrios (neighborhoods)
- Direcciones urbanas asignadas según:
  - Malla vial (road\_network\_addresses)
  - Nomenclatura Barrio - Manzana - Identificación del predio (neighborhood\_block\_houseid\_addresses).

Las entidades y relaciones del modelo de datos definido se explican a continuación (véase figura 19):

- Una comuna puede contener varios barrios.

- Un barrio puede contener varios predios, por ende, direcciones urbanas asignadas según la malla vial o según la nomenclatura Barrio Manzana – Identificación del predio.
- Determinar la relación entre las vías con las demás entidades tiene cierto grado de dificultad, puesto que se debe considerar la ubicación espacial de cada una de estas para encontrar los segmentos de vía que se están ubicados principalmente en las comunas y barrios del municipio de Pasto. Por este motivo, esta entidad fue definida sin relación alguna.

Figura 19. Modelo Entidad - Relación de la base de datos con información del área urbana del municipio de Pasto.



**Migración de datos OSM a PostgreSQL:** Dado que no fue posible encontrar una herramienta que permita migrar los datos recopilados, fue necesaria la codificación de procedimientos parametrizados de migración de datos de los archivos OSM obtenidos a una base de datos relacional con el modelo de datos implementado en el gestor de bases de datos PostgreSQL [35] con extensión de datos espaciales PostGIS [36]. Los procedimientos fueron escritos en el lenguaje de programación Python, los cuales se describen a continuación:

- Un procedimiento de migración de geometrías basadas en polígonos y líneas creado con el fin de migrar los datos de los barrios y las comunas corregidas.
- Un procedimiento de migración basado en puntos creado con el fin de migrar los datos de las direcciones urbanas recopiladas. Dado que cada archivo puede contener direcciones urbanas asignadas según la malla vial o según la nomenclatura barrio – manzana- predio, éste reconoce estos dos tipos de direcciones y los inserta en su respectiva entidad (tabla).

Previamente, cada procedimiento transforma el sistema de referencia espacial nativo de los datos al sistema de referencia espacial WGS84 Web Mercator (EPSG: 3857) [37], esto con el fin de facilitar el manejo de las direcciones urbanas con las herramientas utilizadas para construir el visor cartográfico, las cuales trabajan de forma nativa con este sistema.

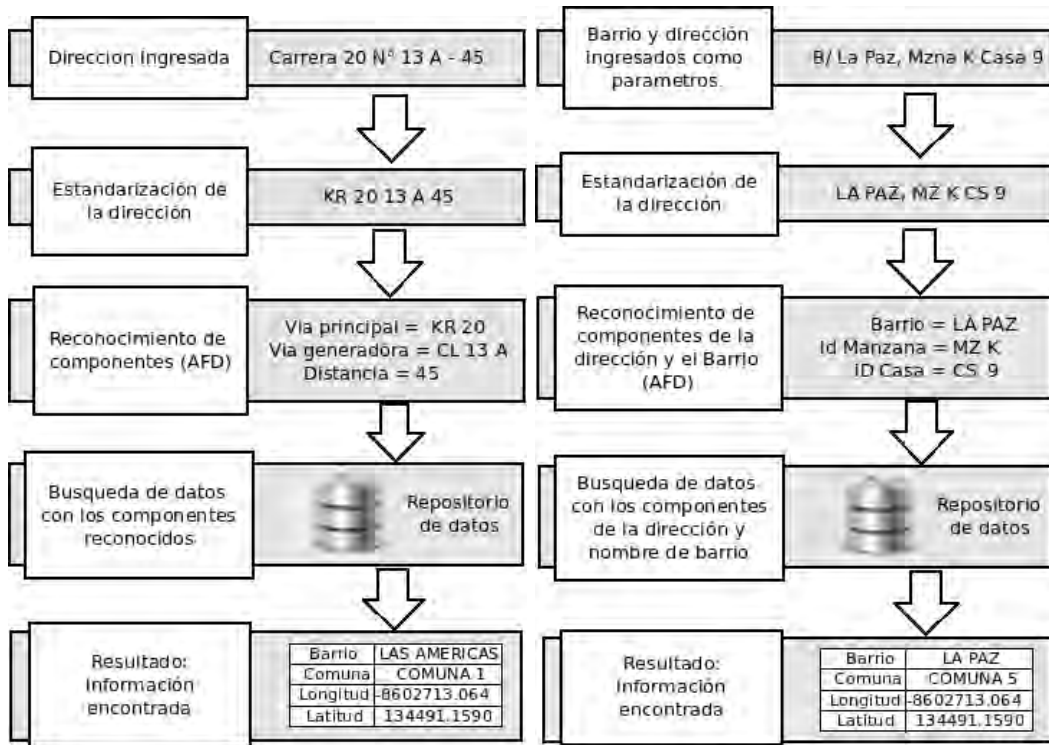
**2.3.2. Construcción del geocodificador de direcciones urbanas.** En esta fase se definieron los procesos de geocodificación de direcciones para el área urbana del municipio de Pasto. Posteriormente, estos fueron implementados en procedimientos almacenados de estandarización y reconocimiento de componentes de las direcciones urbanas para el municipio de Pasto.

### **Definición procesos de geocodificación**

Considerando el proceso de geocodificación definido por [10] y los casos de asignación de direcciones del municipio de Pasto seleccionados para la construcción del geocodificador previamente estudiados, se definieron los procesos de geocodificación de direcciones para direcciones asignadas según la malla vial y según la nomenclatura barrio - manzana – predio. Se tuvo como referencia las normas establecidas en la propuesta de estandarización de direcciones del Instituto Geográfico Agustín Codazzi [25], las cuales fueron adaptadas al contexto del municipio de Pasto.

Los procesos de geocodificación para el área urbana del municipio de Pasto pueden verse en la figura 20.

Figura 20. Procesos de geocodificación para direcciones urbanas del municipio de Pasto



a) Según malla vial

b) Según manzanas

Los Algoritmos 1 y 2 muestran el pseudocódigo de los algoritmos de geocodificación definidos por cada caso de direcciones urbanas.

Algoritmo1 - RoadNetworkGeocoding. Geocodificación de direcciones asignadas según la malla vial.

---

**Algorithm** RoadNetworkGeocoding

---

```

1: procedure GEOCODE(searchedAddress)
2:   find and replace searchedAddress components with standardized nomenclatures using regex
3:   addrElements[]  $\leftarrow$  searchedAddress converted to array using space character as delimiter
4:   state  $\leftarrow$  0
5:   for each element in addrElements do
6:     if state = 0 then
7:       if element is similar to [CL] or [KR] or [DG] then
8:         identify the matching nomenclature
9:         main_road  $\leftarrow$  value of standardized nomenclature identified.
10:        state  $\leftarrow$  1
11:       end if
12:     else if state = 1 then
13:       if element is similar to [0 – 9] or [0 – 9][A – Z] then
14:         main_road  $\leftarrow$  main_road concatenated with the current element identified.
15:         state  $\leftarrow$  2
16:       end if
17:     else if state = 2 then
18:       if element is similar to [A – Z] or BIS or NORTE|SUR|ESTE|OESTE then
19:         main_road  $\leftarrow$  main_road concatenated with the current element identified.
20:         state  $\leftarrow$  2
21:       else if element is similar to [0 – 9] or [0 – 9][A – Z] then
22:         generator_road  $\leftarrow$  value of current element identified.
23:         state  $\leftarrow$  3
24:       end if
25:     else if state = 3 then
26:       if element is similar to [A – Z] or BIS or NORTE|SUR|ESTE|OESTE then
27:         generator_road  $\leftarrow$  generator_road concatenated with the current element identified.
28:         state  $\leftarrow$  3
29:       else if element is similar to [0 – 9] or [0 – 9][A – Z] then
30:         distance  $\leftarrow$  value of current element identified discarding letters.
31:         state  $\leftarrow$  4
32:       end if
33:     end if
34:   end for
35:   if state > 3 then
36:     search required address in database using the recognized address components.
37:     if recognized address components are found then
38:       return standardized address, neighborhood, commune, longitude, latitude found
39:     end if
40:   end if
41: end procedure

```

---

Algoritmo 2 - NBlockHouseidGeocoding. Geocodificación de direcciones asignadas según la nomenclatura Barrio Manzana Predio.

---

**Algorithm** NBlockHouseidGeocoding

---

```

1: procedure GEOCODE(searchedNeighborhood, searchedAddress)
2:   neighborhood  $\leftarrow$  searchedNeighborhoodwithstandardizednomenclatures
3:   address  $\leftarrow$  searchedAddresswithstandardizednomenclatures
4:   addrElements[]  $\leftarrow$  address converted to array using space character as delimiter
5:   state  $\leftarrow$  0
6:   for each element in addrElements do
7:     if state = 0 then
8:       if element is similar to [M Z] then
9:         identify the matching nomenclature
10:        block  $\leftarrow$  value of standardized nomenclature identified.
11:        state  $\leftarrow$  1
12:      end if
13:    else if state = 1 then
14:      if element is similar to [0 - 9] or [A - Z] or [0 - 9][A - Z] then
15:        identify the matching element
16:        block  $\leftarrow$  block concatenated with the current element identified.
17:        state  $\leftarrow$  3
18:      end if
19:    else if state = 2 then
20:      if element is similar to [A - Z] or [CS] then
21:        identify the matching element
22:        houseid  $\leftarrow$  value of current element identified.
23:        state  $\leftarrow$  3
24:      end if
25:    else if state = 3 then
26:      if element is similar to [CS] then
27:        block  $\leftarrow$  block concatenated with the current element identified.
28:        state  $\leftarrow$  3
29:      else if element is similar to [0 - 9] or [0 - 9][A - Z] then
30:        identify the matching element
31:        houseid  $\leftarrow$  'CS' string concatenated with the value of current element identified.
32:        state  $\leftarrow$  4
33:      end if
34:    else if state = 4 then
35:      if element is similar to [A - Z] then
36:        houseid  $\leftarrow$  houseid concatenated with the current element identified.
37:        state  $\leftarrow$  4
38:      end if
39:    end if
40:  end for
41:  if state > 3 then
42:    search required address in database using the recognized address components.
43:    if recognized address components and neighborhood are found then
44:      return standardized address, neighborhood, commune, longitude, latitude found
45:    end if
46:  end if
47: end procedure

```

---

A continuación se describen estos algoritmos:

**a. Ingreso de direcciones:** En este primer paso del algoritmo se definen las características de direcciones urbanas que serán ingresadas y procesadas por el geocodificador. Para esto, se hizo la revisión de las direcciones almacenadas en los registros de delitos del observatorio del delito del municipio de Pasto. Esta revisión sirvió de base para asumir que las direcciones urbanas ingresadas al geocodificador tendrá diversos formatos que no cumple ningún tipo de norma. Algunas de estas direcciones urbanas revisadas pueden observarse en la figura 21.

Figura 21. Algunas direcciones urbanas registradas por el observatorio del delito.

direccion character varying	barrio character varying(75)	direccion character varying	barrio character varying(75)
CALLE 16B No 33-61	MARIDIAZ	MZ 21 CASA 13	NUEVO SOL
CRR 23 No 10-37	SANTIAGO	MZ B1 CASA 4	NUEVA ARANDA
CALLE 19A/3E-20	SANTAFÉ I	MZ 19 CASA 1	VILLA FLOR II
CRA 22F 9 26 OBRERO	SAN JOSE OBRERO	MZ E CASA 7	PANORAMICO I
CRA 6B N 18B-31	SENDOYA	MZ 22 CASA 10	SAN ALBANO
CLL 11 B No 4 14 B /	CHAPAL	MZ 8 CASA 27	LA MINGA
CRA 2 N 19-60	BETANIA	MZ A CASA 22	GUALCALOMA
CaLle 5 Nro. 22 D-42	SAN JOSE OBRERO	MZ 49 CASA 18	EL CHAMBU I
CRA 29 No 19-81	LAS CUADRAS	MZ 51 CASA 22	TAMASAGRA I
CRA 26 N 21B-24	SANTA BARBARA	MZ B CASA B - 13	NUEVO HORIZONTE
KRA 4 E N 21 B 43	SANTA BARBARA	MZ K CASA 167	CAICEDONIA
CRA 5 E N 21 - 25	SANTA BARBARA	MZ 11 CASA 11	LA MINGA
CARRERA 7 N 20 -28	CHILE	MZ 3 CASA 17	NIZA I
CR 7E Nº 19 A 24	BETANIA	MZ 43 CASA 1	VILLA FLOR II
CRA 7 N 19A-40	SANTA FE	MZ D CASA 5	NIÑO JESUS DE PRAGA
Cra 2 19-18	SANTA FE	MZ 19 CASA 26	VILLA FLOR I

a) Según la malla vial

b) Según nomenclatura barrio -  
manzana - predio

A continuación se describen las principales características de las direcciones urbanas revisadas:

- Para el caso de las direcciones urbanas asignadas según malla vial se ingresa la dirección urbana solicitada. Esta debe tener tres componentes: **vial principal, vía generadora y distancia**. Ej. Carrera 20 13 A – 45.
- Para el caso de direcciones urbanas asignadas según la nomenclatura barrio – manzana – predio, deben ingresarse dos datos sobre la dirección urbana solicitada: **el nombre del barrio y la identificación del predio de acuerdo con este caso**. Este último debe tener mínimo dos componentes: **Identificación de la Manzana e Identificación del predio**. Ej. Barrio La Paz, Manzana K - Casa 9.

- Las direcciones ingresadas pueden contener datos adicionales, estos deben ser posteriormente descartados por el procedimiento de normalización de direcciones.

**b. Estandarización de la dirección urbana (líneas 2-3 Algoritmo 1, líneas 2-4 Algoritmo 2):**

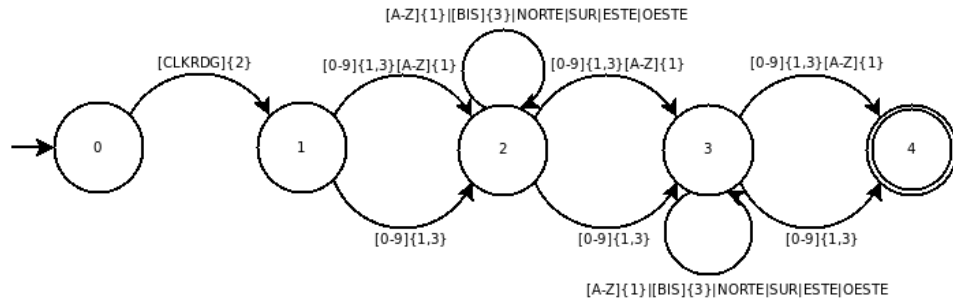
En la etapa de estandarización de direcciones se reemplazan las nomenclaturas y símbolos reconocidos en la dirección urbana ingresada con las sugeridas por la propuesta de estandarización. En la tabla 5 se puede observar algunos ejemplos de reemplazo.

Tabla 5. Casos de reemplazo de símbolos y nomenclaturas en direcciones urbanas ingresadas.

<b>Símbolos/Nomenclaturas reconocidos</b>	<b>Símbolo(s) reemplazo</b>
Texto en minúsculas	Texto en Mayúsculas
CARRERA CRA KRA KR CR CARERA	KR
CALLE CALL CLLE CLL CL CAL CLE CALE	CL
DIAGONAL DIAG DGNAL	DG
MANZANA MZNA MZA MANZ MZN	MZ
CASA CSA CAS  C	CS
N NO NUMERO NUM	Espacio
Signos de puntuación (.,;:-)	Espacio
Caracteres especiales (\$%&/)	Espacio

**c. Reconocimiento de los componentes de la dirección urbana (líneas 5-34 Algoritmo 1, líneas 6-40 Algoritmo 2):** Con el fin de obtener los componentes mínimos de la dirección urbana ingresada y eliminar cualquier tipo de texto adicional, se consideró el geocodificador creado por [12] para implementar autómatas finitos deterministas (AFD), encargados de reconocer los componentes mínimos de cada dirección, descartando o eliminando texto adicional que haya sido registrado junto con la dirección urbana. Para el diseño de estos autómatas se utilizaron expresiones regulares que permiten el reconocimiento de texto a partir de patrones de búsqueda. El AFD definido para las direcciones según la malla vial se muestra en la figura 22.

Figura 22. Autómata Finito Determinista que reconoce los componentes de las direcciones asignadas según la malla vial del municipio de Pasto



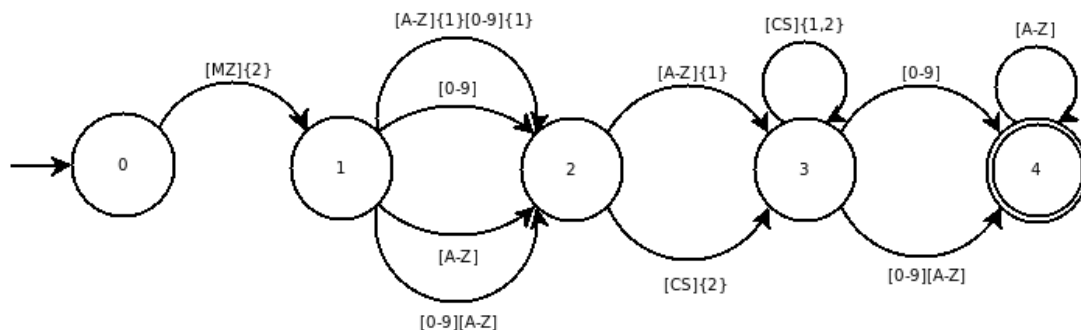
El resultado que arroja el Autómata Finito Determinista se muestra en la tabla 6.

Tabla 6. Resultado del proceso ejecutado por el Autómata Finito Determinista para direcciones urbanas asignadas según la malla vial.

Dirección Ingresada	Dirección estandarizada	Componentes reconocidos (Eliminación de texto Adicional)		
		Vía principal	Vía generadora	Distancia
Carrera 20 #13 a 45	KR 20 13 A 45	KR 20	13 A	45
Calle 9 No 34 – 30 Las Acacias	CL 9 34 30 LAS ACACIAS	CL 9	34	30
Cll 18 26 - 54 Centro	CL 18 26 54 CENTRO	CL 18	26	54
Calle 3 15 62 Caicedo Alto	CL 3 15 62 CAICEDO ALTO	CL 3	15	62

El Autómata Finito Determinista definido para direcciones asignadas según la nomenclatura Barrio-Manzana-Identificación de predio aprovecha el ingreso de los datos Barrio y Dirección urbana por separado, procesando únicamente este último elemento. El dato que corresponde al barrio se procesa mediante una búsqueda aproximada de texto utilizando el algoritmo de similitud de texto Jarowinkler [38]. El procesamiento por separado de estos dos componentes se debe a la identificación de direcciones que no tenían un barrio asociado, dato importante en la geocodificación de este tipo de direcciones. Para este caso, el AFD que procesa el elemento dirección urbana se muestra en la figura 23.

Figura 23. Autómata Finito Determinista que reconoce los componentes de direcciones asignadas según la nomenclatura Barrio-Manzana-Identificación de predio



El resultado que arroja el Autómata Finito Determinista se muestra en la tabla 7.

Tabla 7. Resultado del proceso ejecutado por el Autómata Finito Determinista para direcciones urbanas asignadas según la nomenclatura barrio – manzana – predio.

Dirección Ingresada	Dirección estandarizada	Componentes reconocidos (Eliminación de texto Adicional)	
		Id. Manzana	Id. Predio
Manzana 10 Casa 9	<b>MZ 10 CS 9</b>	MZ 10	CS 9
Mzna 2 Casa 1A	<b>MZ 2 CS 1<sup>a</sup></b>	MZ 9	CS 1A
Barrio La Paz Mz K Cs 9	<b>BARRIO LA PAZ MZ K CS 9</b>	MZ K	CS 9
Mzna 10 F Csa 2 B/ Nueva Aranda	<b>MZ 10 CS 2 B NUEVA ARANDA</b>	MZ 10F	CS 2

**Búsqueda de la información urbana (líneas 35-39 Algoritmo 1, líneas 41-45 Algoritmo 2):** Una vez obtenidos los componentes de la dirección urbana buscada, finalmente se realiza la búsqueda de la dirección en el repositorio de información urbana. Si la dirección es encontrada se retornará los siguiente datos:

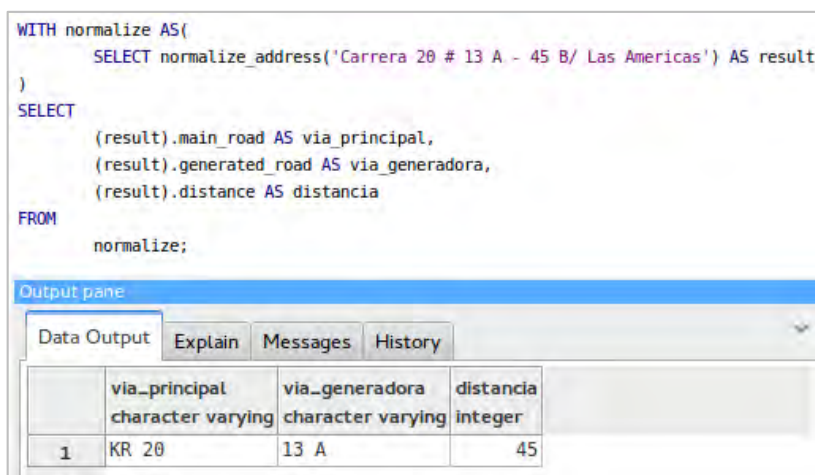
- Dirección urbana estandarizada
- Barrio al que pertenece.
- Comuna a la que pertenece.
- Coordenadas latitud y longitud de ubicación espacial.

**Implementación procesos de geocodificación:** Una vez analizado el geocodificador construido por [2], se optó por implementar los procesos de geocodificación de direcciones en procedimientos almacenados dentro del sistema gestor de bases de datos PostgreSQL con extensión para el soporte de datos espaciales PostGIS. En conjunto, estas herramientas permitieron la manipulación de los datos espaciales mediante recursos de las bases de datos como el lenguaje de consulta estructurado SQL y procedimientos almacenados para el manejo de expresiones regulares y consultas espaciales.

Como resultado de la implementación se tuvo el geocodificador de direcciones urbanas, compuestos por 4 funciones procedurales codificadas en el lenguaje procedural PL/PgSQL de PostgreSQL:

El procedimiento almacenado de estandarización de direcciones asignadas según la malla vial se encarga de estandarizar y reconocer los componentes de las direcciones urbanas del municipio de Pasto. La ejecución del procedimiento almacenado creado para las direcciones se puede ver en la figura 24.

Figura 24. Ejecución del procedimiento de estandarización de direcciones urbanas asignadas según la malla vial



El procedimiento de estandarización de direcciones asignadas según la nomenclatura barrio–manzana – predio se encarga de estandarizar y reconocer los componentes de las direcciones urbanas del municipio de Pasto. La ejecución del procedimiento se puede ver en la figura 25.

Figura 25. Ejecución del procedimiento almacenado de estandarización de direcciones urbanas asignadas según la nomenclatura barrio - manzana - predio

```

WITH normalize AS(
    SELECT normalize_address('Barrio La Paz', 'Manzana K Casa 9') AS result
)
SELECT
    (result).neighborhood AS barrio,
    (result).block AS id_manzana,
    (result).houseid AS id_predio
FROM
    normalize;
    
```

Output pane

Data Output Explain Messages History

	barrio character varying(40)	id_manzana character varying(20)	id_predio character varying(20)
1	LA PAZ	MZ K	CS 9

El procedimiento almacenado de búsqueda de direcciones asignadas según la nomenclatura barrio–manzana – predio se encarga de realizar la búsqueda de la dirección asignada según la malla vial estandarizada mediante sus componentes reconocidos. La ejecución del procedimiento almacenado se puede ver en la figura 26.

Figura 26. Ejecución del procedimiento almacenado de búsqueda de direcciones urbanas asignadas según la malla vial

```

WITH geocode AS (
    SELECT geocode_address('Carrera 20 # 13 A - 45 B\ Las Americas') AS result
)
SELECT
    (result).neighborhood AS barrio,
    (result).address AS direccion,
    (result).commune AS comuna,
    (result).lon AS longitud,
    (result).lat AS latitud
FROM
    geocode;
    
```

Output pane

Data Output Explain Messages History

	barrio character varying(50)	direccion character varying(25)	comuna character varyi	longitud double precision	latitud double precision
1	LAS AMERICAS	KR 20 13 A 45	COMUNA 1	-8602713.06642356	134491.159088312

El procedimiento almacenado de búsqueda de direcciones asignadas según la nomenclatura barrio–manzana – predio se encarga de realizar la búsqueda de la dirección asignada según la nomenclatura barrio - manzana - predio estandarizada, mediante sus componentes reconocidos. La ejecución del procedimiento almacenado se puede ver en la figura 27.

Figura 27. Ejecución del procedimiento almacenado de búsqueda de direcciones urbanas asignadas según la nomenclatura barrio - manzana - predio

```

WITH geocode AS (
    SELECT geocode_address('Br La Paz', 'Manzana K Casa 9') AS result
)
SELECT
    (result).neighborhood AS barrio,
    (result).address AS direccion,
    (result).commune AS comuna,
    (result).lon AS longitud,
    (result).lat AS latitud
FROM
    geocode;

```

	barrio character varying(50)	direccion character varying(25)	comuna character varying(50)	longitud double precision	latitud double precision
1	LA PAZ	MZ K CS 9	COMUNA 4	-8600523.43231094	133201.206467672

**Construcción del visor cartográfico PASTO VIEW:** Una vez construido el geocodificador funcional de direcciones urbana, este fue acoplado con un visor cartográfico denominado PASTO VIEW, cuyo logo inicial se muestra en la figura 28. Este visor permite a cualquier usuario acceder al geocodificador de direcciones urbanas a través de una interfaz gráfica de usuario, ingresando direcciones urbanas y visualizando su correspondiente ubicación urbana de ser encontradas en el repositorio de datos.

Figura 28. Logo visor cartográfico PASTO—VIEW



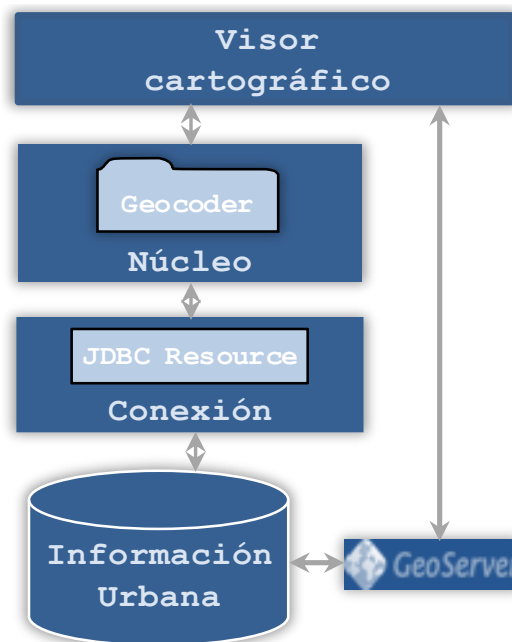
La construcción del visor cartográfico fue realizada en su totalidad bajo el sistema operativo Linux Ubuntu en su versión 14.04. El lenguaje de programación utilizado fue Java, con el entorno de

desarrollo integrado (IDE) Netbeans versión 8.1. El servidor de aplicaciones utilizado fue Glassfish en su versión 3.2.

La arquitectura de PASTO-VIEW se basa en el patrón Modelo Vista Controlador (MVC), implementado en tres módulos: el modulo visor, el modulo núcleo y el modulo conexión a base de datos. La figura 29 muestra la arquitectura del visor cartográfico.

Los requerimientos y el análisis UML definidos para el visor cartográfico PASTO-VIEW se pueden consultar en el Anexo D.

Figura 29. Arquitectura del visor cartográfico PASTO-VIEW



- **Modulo Visor Cartográfico:** Permite el ingreso de direcciones urbanas para visualizar su ubicación urbana, así como también la carga y procesamiento de archivos separados por comas CSV. Este módulo se conecta con el modulo núcleo de forma directa.

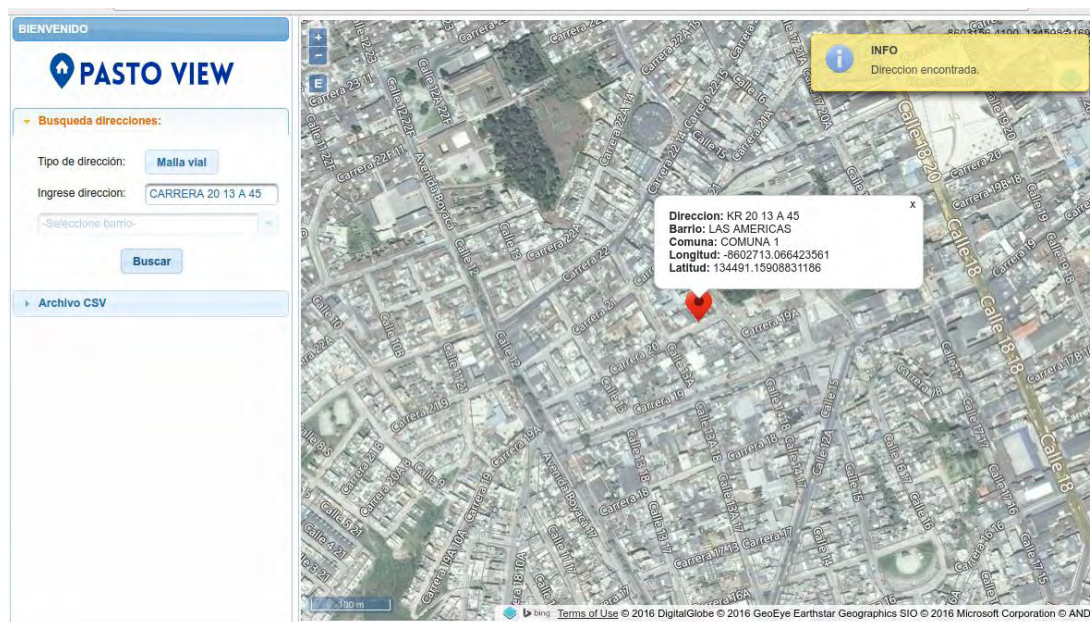
Específicamente este módulo se compone de una página con extensión xhtml del framework Java Server Faces (JSF), encargada de interactuar con el usuario e intercambiar información con el controlador que hace parte del kernel de PASTO VIEW. Para la creación de la interfaz gráfica fue utilizado el framework Primefaces en su versión 3.4. Para la visualización de mapas fue utilizada la interfaz programable de aplicaciones (API) OpenLayers en su versión 3.14, la cual está escrita en el lenguaje de programación JavaScript. Esta interfaz se conecta con

servicios como Bing Maps y Open Street Maps para la carga de mapas de referencia. A continuación, se describe la página xhtml de acceso principal dentro de este módulo:

- index.xhtml: Mediante esta página el usuario accede a dos funcionalidades: La consulta y visualización de la ubicación de direcciones urbanas en el área urbana del municipio de Pasto y la geocodificación de datos contenidos en archivos separados por comas CSV.

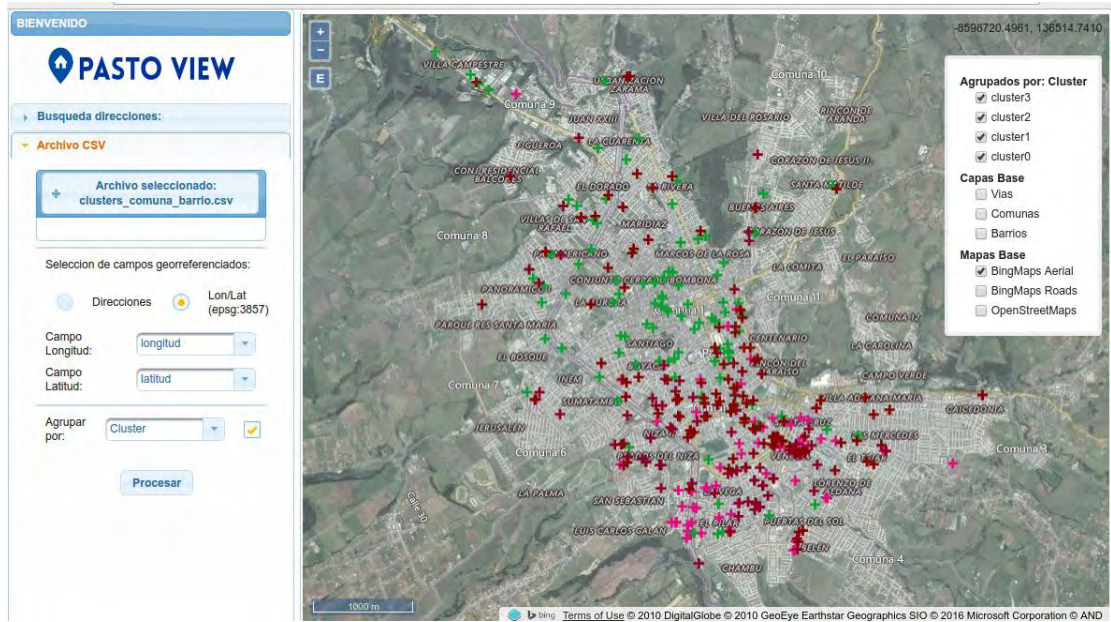
La búsqueda de direcciones puede realizarse según los dos casos de asignación de direcciones seleccionados para esta investigación: según malla vial ò según la nomenclatura barrio - manzana predio. La figura 30 muestra la implementación de esta búsqueda.

Figura 30. Búsqueda de direcciones, funcionalidad implementada en PASTO VIEW



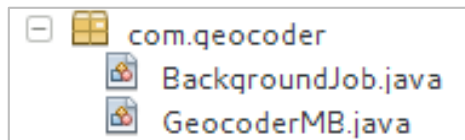
El procesamiento de archivos CSV se realiza mediante la selección de campos georreferenciados como pueden ser direcciones urbanas o coordenadas latitud/longitud las cuales deben estar procesados con el sistema de referencia espacial EPSG:3857. De igual forma se dispone de una opción que permite agrupar datos según el campo que se haya seleccionado. Finalmente, el usuario podrá consultar mediante un dialogo de información las características o campos restantes asociados a los puntos desplegados haciendo clic sobre cada uno de estos. La figura 31 muestra la implementación de este procesamiento.

Figura 31. Procesamiento de archivo CSV, datos agrupados por el atributo Cluster.  
Funcionalidad implementada en PASTO VIEW



- **Módulo Núcleo:** En este módulo se encuentra un único paquete encargado de procesar la información solicitada por el usuario, el cual se describe a continuación:
  - Paquete Geocoder: En este paquete se almacenan una única clase encargada de procesar los datos asociados con la búsqueda de direcciones urbanas que son solicitados por el usuario. La estructura de este paquete puede verse en la figura 32.

Figura 32. Paquete geocoder visor PASTO VIEW



BackgroundJob.java es la clase encargada de eliminar los datos temporales almacenados en la base de datos generados por los usuarios que acceden al sistema en el procesamiento de archivos CSV. El único método de esta clase denominado hoursJob() se ejecuta cada hora eliminando aquellos datos cuyo timestamp (fecha y hora) de inserción sea superior a una hora en comparación con el timestamp que corresponde al momento de la ejecución del procedimiento.

GeocoderMB.java es la clase encargada de conectarse con el geocodificador de direcciones urbanas para la búsqueda de direcciones urbanas y el procesamiento de los archivos CSV. Esta clase tiene dos métodos principales: Uno de ellos se denomina processAddress() el cual genera y ejecuta una sentencia SQL sobre el geocodificador de direcciones solicitando los datos asociados a la dirección buscada para ser recuperados y visualizados posteriormente. El procedimiento de ejecución se puede observar en la figura 33.

Figura 33. Código de conexión del visor PASTO VIEW con el geocodificador de direcciones urbanas.

```

sql = ""
+ "WITH geocoded AS(\n"
+ "  SELECT\n"
+ "    geocode_address('" + searchedRAddress + "') AS result\n"
+ ")\n"
+ "SELECT\n"
+ "  (result).address,\n"
+ "  (result).neighborhood,\n"
+ "  (result).commune,\n"
+ "  (result).lon,\n"
+ "  (result).lat\n"
+ "FROM\n"
+ "  geocoded\n"
+ "WHERE\n"
+ "  (result).lon IS NOT NULL AND (result).lat IS NOT NULL;";

try {
  ResultSet rs = connectionJdbcMB.consult(sql);
  if (rs.next()) {
    resultAddress = rs.getString("address");
    resultNeighborhood = rs.getString("neighborhood");
    resultCommune = rs.getString("commune");
    resultLongitude = rs.getDouble("lon");
    resultLatitude = rs.getDouble("lat");
    JsfUtil.addSuccessMessage("Dirección encontrada");
  } else {
    resultLongitude = 0;
    resultLatitude = 0;
    JsfUtil.addErrorMessage("Dirección no encontrada.");
  }
} catch (SQLException e) {
  System.out.println("Error 1 en " + this.getClass().getName() + ":" + e.toString());
}

```

El método processCSV() se encarga de leer archivos CVS, identificando el nombre de los campos y procesando cada registro almacenando en dicho archivo de acuerdo con la configuración seleccionada por el usuario.

- **Conexión a base de datos:** Para conectar PASTOVIEW con el repositorio de datos, se utiliza la anotación @Resource disponible con la edición empresarial de Java J2EE, esta permite referenciar y utilizar una conexión JDBC creada desde el servidor de aplicaciones Glassfish.

**2.3.3. Construcción visor cartográfico de eventos delictivos.** En esta fase de la investigación el geocodificador fue acoplado con SIGEODEP, el sistema georreferenciado del observatorio del delito del municipio de Pasto mediante la construcción del visor cartográfico SIGEODEP-SIG, permitiendo visualizar la geocodificación de los eventos delictivos fatales y no fatales registrados por este ente gubernamental a nivel de direcciones urbanas.

**Geocodificación de eventos delictivos fatales y no fatales:** la base de datos de delitos del sistema SIGEODEP se compone de 149 tablas, entre las cuales se identifican dos principales: fatal\_injuries (lesiones fatales) y non\_fatal\_injuries (lesiones no fatales), que contienen los principales atributos de los eventos delictivos que referencian a otras tablas o son referenciados mediante claves foráneas. La figura 34 muestra la estructura de estas tablas.

Figura 34. Tablas base de datos SIGEODEP que almacenan los eventos delictivos.

non_fatal_injuries	fatal_injuries
non_fatal_injury_id: INTEGER [ PK ]	fatal_injury_id: INTEGER [ PK ]
injury_id: SMALLINT	injury_id: SMALLINT
checkup_date: DATE	injury_date: DATE
checkup_time: TIME	injury_time: TIME
injury_date: DATE	injury_address: VARCHAR
injury_time: TIME	injury_neighborhood_id: INTEGER
injury_address: VARCHAR	injury_place_id: SMALLINT
injury_neighborhood_id: INTEGER	activity_id: SMALLINT
injury_place_id: SMALLINT	intentionality_id: SMALLINT
activity_id: SMALLINT	use_alcohol_id: SMALLINT
intentionality_id: SMALLINT	use_drugs_id: SMALLINT
use_alcohol_id: SMALLINT	burn_injury_degree: SMALLINT
use_drugs_id: SMALLINT	burn_injury_percentage: SMALLINT
burn_injury_degree: SMALLINT	submitted_patient: BOOLEAN
burn_injury_percentage: SMALLINT	destination_patient_id: SMALLINT
submitted_patient: BOOLEAN	input_timestamp: TIMESTAMP
destination_patient_id: SMALLINT	injury_day_of_week: VARCHAR
input_timestamp: TIMESTAMP	victim_id: INTEGER
health_professional_id: SMALLINT	alcohol_level_victim_id: SMALLINT
non_fatal_data_source_id: SMALLINT	alcohol_level_victim: SMALLINT
mechanism_id: SMALLINT	code: VARCHAR
user_id: SMALLINT	area_id: SMALLINT
injury_day_of_week: VARCHAR	victim_place_of_origin: VARCHAR
victim_id: INTEGER	id_state_date: SMALLINT
submitted_data_source_id: SMALLINT	id_state_time: SMALLINT
id_state_checkup_date: SMALLINT	quadrant_id: SMALLINT
id_state_checkup_time: SMALLINT	
id_state_injury_date: SMALLINT	
id_state_injury_time: SMALLINT	
quadrant_id: SMALLINT	

Para geocodificar los eventos delictivos dispuestos, fue seleccionado el campo dirección (injury\_id) como parámetro de las funciones de geocodificación para asignar a cada dirección encontrada sus correspondientes coordenadas geográficas. Cabe mencionar que muchas de los datos encontrados no correspondían a direcciones urbanas o no tenían los elementos mínimos para ser procesados, por lo cual hubo un gran número de eventos delictivos que fue imposible geocodificar. Algunos ejemplos de estos datos pueden verse en la figura 35.

Figura 35. Algunos ejemplos de datos que no corresponden a direcciones urbanas, registrados en el campo injury\_id.

injury_address character varying(300)
<NULL>
ESQUINA IGLESIA
ANGANROY
FATIMA
VIA
FUNDACION EZEQUIEL MORENO DIAZ
VIA
VIA
COL. CHAMPAGNATH
LA MISMA
CL 16 Y 15
POR LA PRINCIPAL
LA MISMA
TAMASAGRA
FRENTE CLINICA PALERMO
SIN DATO

La ejecución de los procesos de geocodificación de direcciones urbanas dio como resultado la creación de 2 nuevas tablas dentro del esquema de base de datos existente: geocoded\_fatal\_injuries (eventos delictivos fatales con registros geocodificados) y geocoded\_non\_fatal\_injuries (eventos delictivos no fatales con registros geocodificados), los cuales serán referenciadas a partir del identificador único de cada delito.

La creación de estas tablas con los procedimientos SQL utilizados se pueden observar en el Anexo E.

**Construcción del visor cartográfico SIGEODEP SIG:** Fue construido el visor cartográfico denominado SIGEODEP SIG, cuyo logo inicial se muestra en la figura 36. Este sistema permite visualizar el resultado de los procesos de geocodificación de direcciones urbanas ejecutados sobre de los eventos delictivos fatales y no fatales registrados por el observatorio del delito del municipio de Pasto en mapas de puntos y de calor. Para soportar funcionalidades como el cruce de variables y gestión de usuarios fueron reutilizados algunos componentes implementados en el sistema georreferenciado SIGEODEP. De igual forma fue reutilizada la funcionalidad de procesamiento de archivos CSV implementada en PASTO VIEW.

Figura 36. Logo SIGEODEP SIG

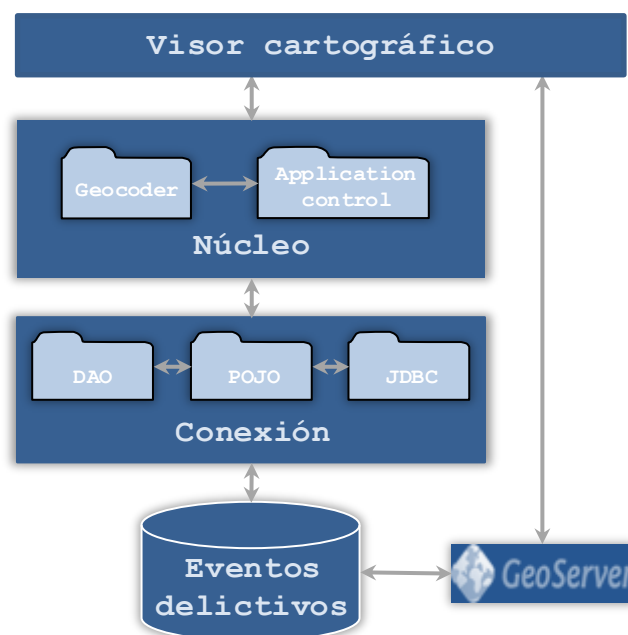


La construcción del visor cartográfico fue realizada en su totalidad bajo el sistema operativo Linux Ubuntu en su versión 14.04. El lenguaje de programación utilizado fue Java, con el entorno de desarrollo integrado (IDE) Netbeans versión 8.1. El servidor de aplicaciones utilizado fue Glassfish en su versión 3.2.

La arquitectura de SIGEODEP SIG se basa en el patrón Modelo Vista Controlador (MVC), implementado en tres módulos: el modulo visor, el modulo núcleo y el modulo conexión a la base de datos. La publicación de mapas y datos espaciales es soportada por GeoServer [39] un servidor de código abierto que se encarga de la gestión y publicación óptima de mapas temáticos. La figura 37 muestra la arquitectura del visor cartográfico.

Los requerimientos y el análisis UML definidos para el visor cartográfico SIGEODEP SIG se pueden consultarse en el Anexo G.

Figura 37. Arquitectura SIGEODEP SIG

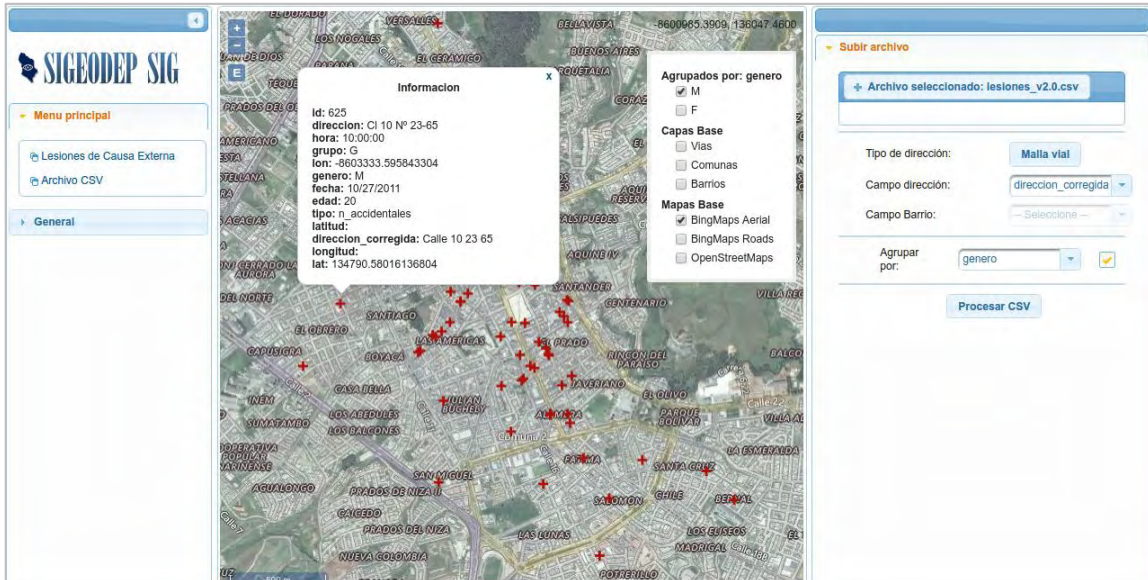


- **Modulo Visor cartográfico:** Modulo que permite al usuario cargar, visualizar e interactuar con los mapas del área urbana del municipio de Pasto, implementando funcionalidades como la creación dinámica de mapas de puntos y de calor permitiendo el análisis espacial de eventos delictivos a nivel de direcciones urbanas mediante el cruce de variables. De igual forma ese modulo cuenta con la funcionalidad de procesamiento de archivos separados por comas CSV para visualizar la ubicación de cada evento o registro encontrado permitiendo agrupar estos datos de acuerdo con una característica o campo determinado. Este módulo se conecta con el modulo núcleo y el módulo gestión de mapas de forma directa.

Específicamente, este módulo se compone de páginas xhtml del framework Java Server Faces (JSF) las cuales se encargan de interactuar con el usuario e intercambiar información con los controladores que hacen parte del kernel de SIGEODEP SIG. Se identifican frameworks como Primefaces en su versión 3.4 para la creación de la interfaz gráfica, y Highcharts en su versión 4.12 para la creación dinámica de gráficos estadísticos. Para la visualización de mapas fue utilizada la interfaz programable de aplicaciones (API) denominada OpenLayers en su versión 3.14. Esta interfaz se conecta con servicios como Bing Maps y Open Street Maps para la carga de mapas de referencia, con el módulo de gestión de Mapas (GeoServer) y permite la interacción espacial con los mapas de puntos y de calor generados. A continuación, se describen las principales páginas que forman el módulo de interfaz gráfica

csvData.xhtml: Esta página permite al usuario cargar y visualizar a nivel de direcciones urbanas los datos almacenados en archivos separados por comas CSV. Se debe especificar los campos del archivo donde se encuentran estos datos georreferenciados según los casos de direcciones urbanas soportados por el geocodificador. La figura 38 muestra la implementación de esta funcionalidad.

Figura 38. Procesamiento y visualización de eventos almacenados en un archivo CSV a nivel de direcciones urbanas.



Injuries.xhtml: Esta página permite al usuario visualizar eventos delictivos fatales y no fatales procesados a nivel de direcciones urbanas generando mapas de puntos y mapas de calor mediante el previo cruce de variables. En el caso de mapas de puntos (véase figura 39) el usuario puede navegar sobre los mapas generados y visualizar datos relacionados con cada evento delictivo georreferenciado. En el caso de los mapas de calor generados (véase figura 40) el usuario puede seleccionar áreas de interés donde se visualicen zonas o patrones visuales de interés y generar gráficos estadísticos sobre el conteo de delitos presentes en esa zona.

Figura 39. Visualización de mapas de puntos procesados a nivel de direcciones urbanas

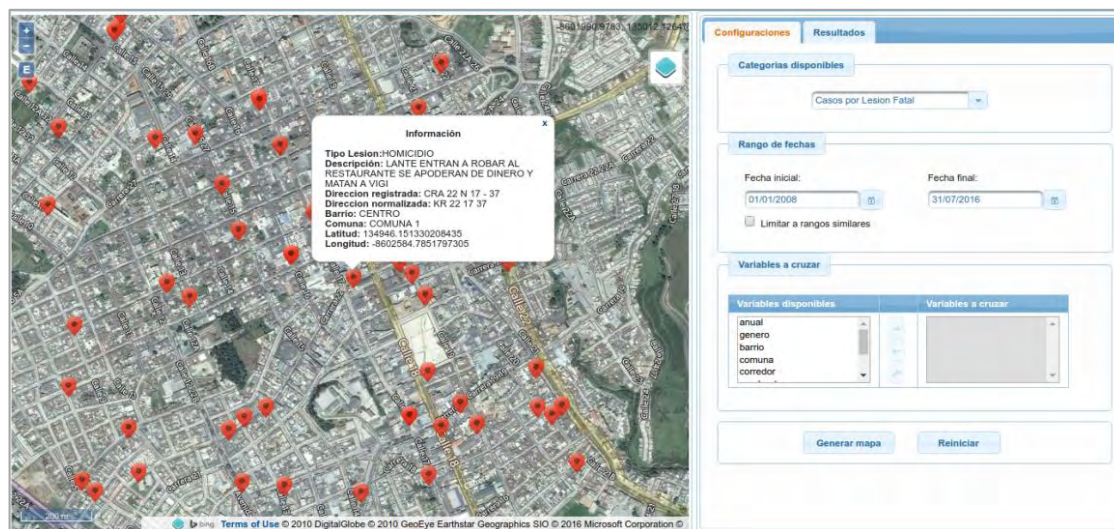
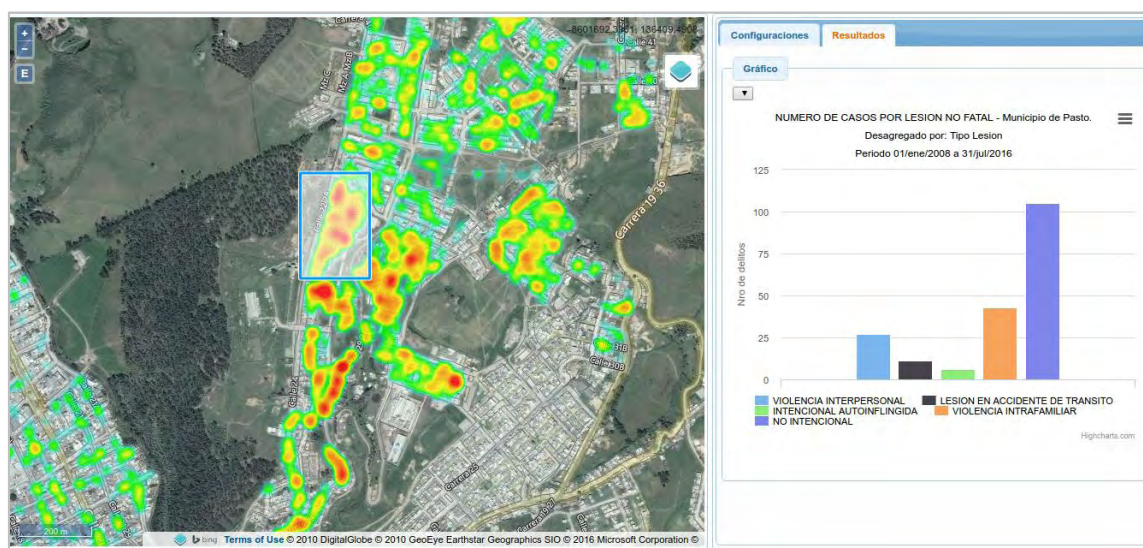


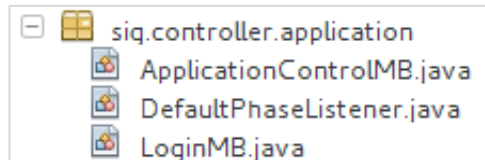
Figura 40. Visualización de mapas de calor procesados a nivel de direcciones urbanas



- **Módulo kernel:** En este módulo se reúnen los paquetes encargados de procesar la información solicitada por el usuario, como también de controlar las sesiones de usuarios invitados. A continuación se describen los paquetes que componen el kernel del sistema:
  - **Application:** En este paquete se almacenan las clases encargadas de gestionar los usuarios invitados que ingresan al visor cartográfico. Dado que funcionalidades como el análisis de

variables ejecutan procesos que interactúan directamente con la base de datos, el manejo de sesiones permite controlar la información generada por cada usuario en su interacción con las funcionalidades del visor. Este módulo está conformado por las clases mostradas en la Figura 41.

Figura 41. Paquete Application Control



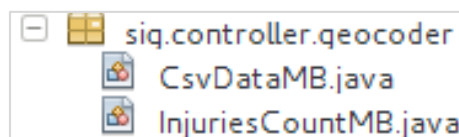
ApplicationControlMB.java es la clase encargada de agregar, consultar, configurar y eliminar sesiones por cada usuario que ingresa al visor cartográfico.

LoginMB.java es la clase encargada de inicializar variables necesarias que usa el sistema, además de configurar atributos que identifican a cada usuario que ingresa a la aplicación.

DefaultPhaseListener.java es la clase que se ejecuta cada vez que se envía una petición al servidor de aplicaciones, verificando que el tiempo de sesión haya caducado. De darse este caso, el sistema elimina todas las configuraciones asociadas al usuario.

- **Geocoder:** En este paquete se almacenan las clases encargadas de procesar los datos solicitados por el usuario invitado relacionados con el procesamiento de archivos CSV y el procesamiento de cruce de variables para su posterior visualización. Este módulo está conformado por las clases mostradas en la Figura 42.

Figura 42. Paquete Geocoder



CsvDataMB.java es la clase encargada de ejecutar la búsqueda de direcciones urbanas y el procesamiento de los archivos CSV. El método principal denominado processCSV() se encarga de leer archivos CVS, identificando el nombre de los campos y procesando cada registro almacenando en dicho archivo de acuerdo con la configuración seleccionada por el usuario.

InjuriesCountMB.java es la clase encargada de realizar todo el proceso de captura, procesamiento y preparación de los resultados asociados al cruce de variables configurado por el usuario. Los principales métodos implementados en esta clase se describen a continuación.

processAddressCountInjuries(): Método principal encargado que integra los métodos necesarios para realizar el proceso de cruce de variables.

createIndicatorConsult(): Método encargado de construir la sentencia SQL que realiza el cruce de las variables seleccionadas por el usuario considerando los valores configurados.

saveIndicatorAddresses(): Método encargado de recorrer los campos generados en el resultado de la consulta ejecutada insertando cada registros encontrado dentro de la tabla denominada indicators\_addresses en la base de datos de delitos.

createCombinations(): Método encargado de crear las posibles combinaciones entre los valores de las variables que fueron seleccionadas por el usuario, insertándolas en la tabla indicators\_addresses.

groupingValues(): Método encargado de agrupar los registros encontrados en el método saveIndicatorsAddresses() y realizar el conteo de estos de acuerdo a cada combinación generada en el método createCombinations().

removeUnusedAddressCombinations(): Metodo encargado de eliminar los registros encontrados en el métodos saveIndicatorsAddresses() que no se encuentren dentro de las combinaciones generadas en el método createCombinations().

checkValidPoints(): Método encargado de encontrar dentro de los registros encontrados en el métodos saveIndicatorsAddresses() aquellos que tienen asignadas coordenadas geográficas latitud/longitud.

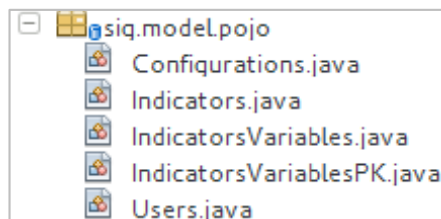
createGeoserverParameters(): Método encargado de configurar las variables necesarias para solicitar a Geoserver la publicación del mapa seleccionado por el usuario.

- **Modulo conexión:** En este módulo se reúnen los paquetes encargados de gestionar la conexión con la base de datos del observatorio del delito. Para esta conexión fueron utilizados recursos

como la API Java Persistence Application (por sus siglas, JPA), Java Data Base Connector (por sus siglas, JDBC). A continuación se describen los paquetes que componen a este módulo.

- **Paquete POJO:** Contiene el modelo de la base de datos representado en entidades, en otras palabras son entidades mapeadas de la base de datos, son las clases de la base de datos con iguales atributos, todo esto lo hace el framework de gestión de la persistencia de los datos JPA (java Persistence API). Las clases que componen a este paquete se muestran en la figura 43.

Figura 43. Clases del paquete POJO



Configurations.java mapea la tabla de la base de datos llamada Configurations. En esta tabla se almacena los valores de configuración de la aplicación como el usuario y contraseña de acceso a la base de datos-

Indicators.java mapea la tabla de la base de datos llamada Indicators. En esta tabla se almacenan la información correspondiente a las categorías de eventos delictivos y sus variables correspondientes.

IndicatorsVariables.java e IndicatorsVariablesPK.java mapean la tabla de la base de datos llamada Indicators\_Variables. En esta tabla se almacenan los posibles valores que pueden tomar las variables cargadas.

Users.java mapean la tabla de la base de datos llamada Users. En esta tabla se almacenan las configuraciones de los usuarios invitados que han ingresado al visor cartográfico.

- **Paquete DAO:** Se compone por un conjunto de clases que se encargan de realizar el acceso a la información contenida en la base de datos, cuando se desee acceder a la base de datos, ya sea para actualizar, borrar, insertar y/o crear nuevos datos, necesariamente se hace uso de este paquete (Véase figura 44).

AbstractFacade.java: es la clase padre que heredar  a las dem s clases todos sus m todos, este proceso de la herencia se encarga de realizarlo el framework llamado Java Persistence API (JPA).

ConfigurationsFacade.java: permite el acceso CRUD a los datos de la entidad Configurations. Esta entidad corresponde a la tabla Configurations de la base de datos.

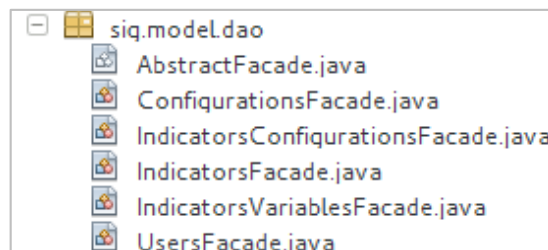
IndicatorsFacade.java: permite el acceso CRUD a los datos de la entidad Indicators. Esta entidad corresponde a la tabla Indicators de la base de datos..

IndicatorsConfigurationsFacade.java: permite el acceso CRUD a los datos de la entidad IndicatorsConfigurations. Esta entidad corresponde a la tabla Indicators de la base de datos.

IndicatorsVariables.java: permite el acceso CRUD a los datos de la entidad IndicatorsVariables. Esta entidad corresponde a la tabla IndicatorsVariables de la base de datos.

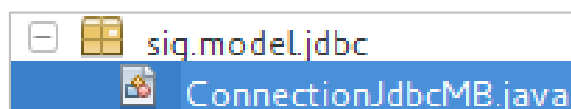
UsersFacade.java: permite el acceso CRUD a los datos de la entidad Users. Esta entidad corresponde a la tabla Users de la base de datos, por ende lleva su mismo nombre.

Figura 44. Clases del Paquete DAO



- **Paquete Jdbc:** Se compone de una  nica clase encargada de ejecutar sentencias SQL creadas en la clase InjuriesCountMB y recuperar los registros retornados mediante la API Java DataBase Connection – JDBC (v ase figura 45). Esta configuraci n de conexi n de datos fue reutilizada del sistema georreferenciado SIGEODEP.

Figura 45. Clases del Paquete JDBC



- **Esquemas de bases de datos.** La base de datos del observatorio del delito tendrá instalados dos esquemas de datos:
  - El esquema public donde se almacenará el esquema que la base de datos del sistema SIGEODEP. El diseño de la base de datos fue realizado en investigación realizada por el grupo de investigación GRIAS de la Universidad de Nariño [3].
  - El esquema denominado geocoder donde se almacenará la información urbana recopilada en las fases iniciales de este proyecto de investigación.

**2.3.4. Pruebas de funcionamiento.** A continuación se describen las pruebas ejecutadas al geocodificador de direcciones urbanas del municipio de Pasto construido que permitieron evaluar aspectos del procesamiento de direcciones urbanas en diferentes casos de prueba.

**Casos de prueba definidos:** Los casos de prueba definidos para probar el funcionamiento del geocodificador de direcciones urbanas son los siguientes:

- Caso de prueba 1. Estandarización de direcciones asignadas según la malla vial.
- Caso de prueba 2. Estandarización de direcciones asignadas según la nomenclatura Barrio - Manzana – Predio.
- Caso de prueba 3. Geocodificación por comuna de direcciones urbanas asignadas según la malla vial.
- Caso de prueba 4. Geocodificación de direcciones urbanas ubicadas en la comuna 3. Caso nomenclaturas Barrio Manzana Predio.

**Ejecución de casos de prueba:** Como datos de prueba se seleccionaron direcciones urbanas registradas por el observatorio del delito del municipio de Pasto que contenían los elementos mínimos requeridos por cada caso de direcciones urbanas implementado en el geocodificador. Fueron eliminados registros duplicados. Fueron seleccionadas 600 direcciones registradas en cada comuna del municipio de Pasto

**Caso de prueba 1. Estandarización direcciones asignadas según la malla vial:** El objetivo de esta prueba fue verificar los resultados arrojados por el procedimiento almacenado de normalización en cada caso de dirección urbana sin formato ingresada. Se seleccionaron 16

direcciones urbanas con formatos complejos de normalizar. La métrica de evaluación se califica con los valores Correcto ò Incorrecto. Los resultados pueden observarse en la tabla 8 y figura 46.

Tabla 8. Resultados individuales caso de prueba para el procedimiento almacenado de Estandarización de direcciones urbanas.

<b>Dirección de prueba</b>	<b>Resultado</b>	<b>Calific.</b>
Calle 12 #4-41 B/ Chapal	CL 12 4 41	Correcto
CRA 22 B 10 SUR 19 SAN MIGUEL DE ALTAMIRA	KR 22 B 10 SUR 19	Correcto
CALLE 12C BIS N 4A-18	CL 12 C BIS 4 A 18	Correcto
DIAGONAL 16 N° 14E/27	DG 16 14 E 27	Correcto
CRA 45 #18-30 EDIF. ALICANTE	KR 45 18 30	Correcto
12 DE OCTUBRE DIAG 15 A NO 2 45	DG 15 A 2 45	Correcto
2A ETAPA CRA 1B #12 – 14	KR 1 B 12 14	Correcto
CARRERA 13 N° 16-31, APTO 501	KR 13 16 31	Correcto
CARRERA 14 A No 1 B 49 B/ CAICEDO	KR 14 1 B 49	Correcto
CRA 36 A CALLE 7 OESTE N 129	KR 36 A 7 OESTE 129	Correcto
CRA 3 A N 11 C 74	KR 3 A 11 C 74	Correcto
Cra 3 N 21B-102B/ STA BARBARA	KR 3 21 B 102	Correcto
Cra 40 A NUMERO 18A 53 AV PANAM.	KR 40 A 18 A 53	Correcto
CRA. 45 CL. 15 D No. 69 <sup>a</sup>	KR 45 15 D 69	Correcto
CRA 4 DGL 16A 15 B/ MIRAFLORES	KR 4 16 A 15	Correcto
SANTA ISABEL CALLE 5 SUR 24 – 58	CL 5 SUR 24 58	Correcto

Figura 46. Resultados Caso de prueba 1 para el procedimiento almacenado de estandarización de direcciones urbanas



**Caso de prueba 2. Estandarización de direcciones asignadas según la nomenclatura Barrio - Manzana - Predio:** El objetivo de esta prueba fue verificar los resultados arrojados por el procedimiento almacenado de normalización en cada caso de dirección urbana sin formato ingresada. Se seleccionaron 16 direcciones urbanas entre las direcciones urbanas de prueba con formatos complejos de normalizar. La métrica de evaluación se califica con los valores Correcto ò Incorrecto. Los resultados pueden observarse en la tabla 9 y figura 47.

Tabla 9. Resultados individuales Caso de Prueba 2

Barrio ingresado	Dirección ingresada	Resultado	Calific.
PANORAMICO	MANZAN E CASA 8	PANORAMICO, MZ E CS 8	Correcto
NUEVA ARANDA	MANZ. B6 CASA 3 NUEVA ARANDA	NUEVA ARANDA, MZ B6 CS 3	Correcto
BR NUEVA ARANDA	"MZ II CASA 14 TERCERA ETAPA NUEVA ARANDA"	NUEVA ARANDA, MZ 11 CS 14	Correcto
SAN CARLOS	"B S AN CAROLOS MZNA O CASA 17"	SAN CARLOS, MZ O CS 17	Correcto

SAN ALBANO	SAN ALBANO MZNA 24 C 2	SAN ALBANO, MZ 24 CS 2	Correcto
SANTA MONICA	MANZANA C CASA 38A BARRIO SANTA MONICA PASTO	SANTA MONICA, MZ C CS 38A	Correcto
ARNULFO GUERRERO	ARNULFO GUERRERO MZNA 12 CASA 5	ARNULFO GUERRERO, MZ 12 CS 5	Correcto
BR LA MINGA	B/LA MINGA MZ 21 CASA 32	LA MINGA, MZ 21 CS 32	Correcto
EL RINCON DE PASTO	"RINCON DE PASTO MZNA E CASA 7"	EL RINCON DE PASTO, MZ E CS 7	Correcto
LAS LAJAS	BR VILLA LAS LAJAS MZ F CS 10	LAS LAJAS, MZ F CS 10	Correcto
GUALCALOM A	GUALCALOMA IV MZ I CASA 12	GUALCALOMA, MZ I CS 2	Correcto
VILLA FLOR II	VILLA FLOR 2, MZ 31 CASA 21	VILLA FLOR, MZ 31 CS 21	Correcto
"SAN ALBANO"	MANZANA 5 CASA7	SAN ALBANO, MZ 5 CS 7	Correcto
EL RINCON DE PASTO	MAZ G CASA 24	EL RINCON DE PASTO, MZ G CS 24	Correcto
Nueva aranda	manzana F5 casa 3	NUEVA ARANDA, MZ F5 CS 3	Correcto
BARRIO GRANADA	"MAZ O CASA 15 4° ETAPA"	GRANADA, MZ O CS 15	Correcto
NUEVA ARANDA	MANZANA E3 CASA 22 NUEVA ARANDA COMFAMILIAR	NUEVA ARANDA, MZ E3 CS 22	Correcto

Figura 47. Resultado caso de prueba 2 Procedimiento almacenado de estandarización

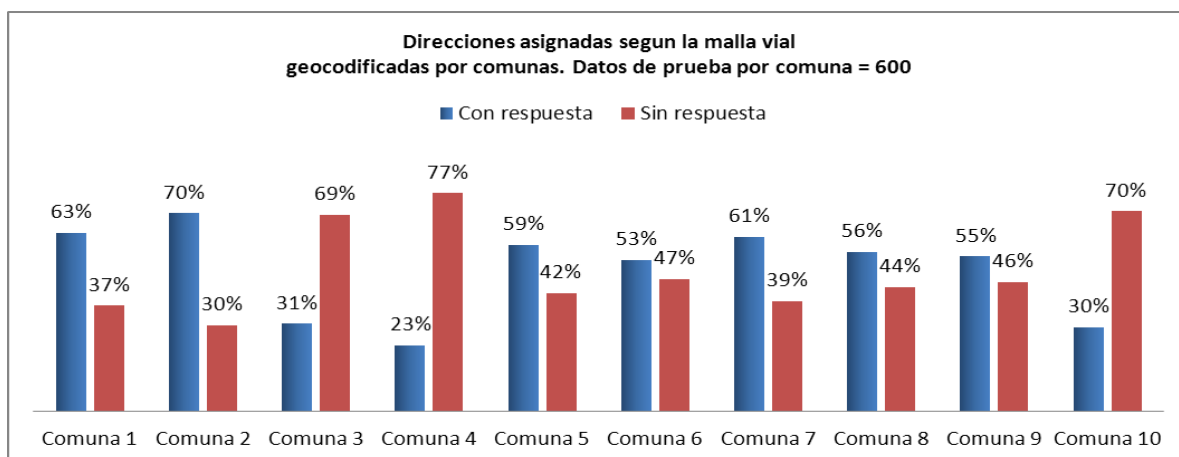


**Caso de prueba 3. Geocodificación por comuna de direcciones urbanas asignadas según la malla vial:** El objetivo de esta prueba fue verificar los resultados arrojados por procedimiento almacenado de geocodificación al ejecutarse con direcciones urbanas registradas en las comunas 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10. Por cada comuna se seleccionaron 600 datos de direcciones urbanas que tiene con los componentes mínimos de direcciones urbanas asignadas según la malla vial sin registros duplicados. Los resultados de la prueba pueden observarse en la tabla 10 y figura 48.

Tabla 10. Resultados de geocodificación de direcciones ejecutado por comuna: direcciones asignadas según la malla vial.

	Con respuesta		Sin respuesta	
<b>Comuna 1</b>	377	63%	223	37%
<b>Comuna 2</b>	418	70%	182	30%
<b>Comuna 3</b>	185	31%	415	69%
<b>Comuna 4</b>	140	23%	460	77%
<b>Comuna 5</b>	351	59%	249	42%
<b>Comuna 6</b>	320	53%	280	47%
<b>Comuna 7</b>	368	61%	232	39%
<b>Comuna 8</b>	337	56%	263	44%
<b>Comuna 9</b>	327	55%	273	46%
<b>Comuna 10</b>	178	30%	422	70%

Figura 48. Resultado geocodificación de direcciones de prueba por comuna.

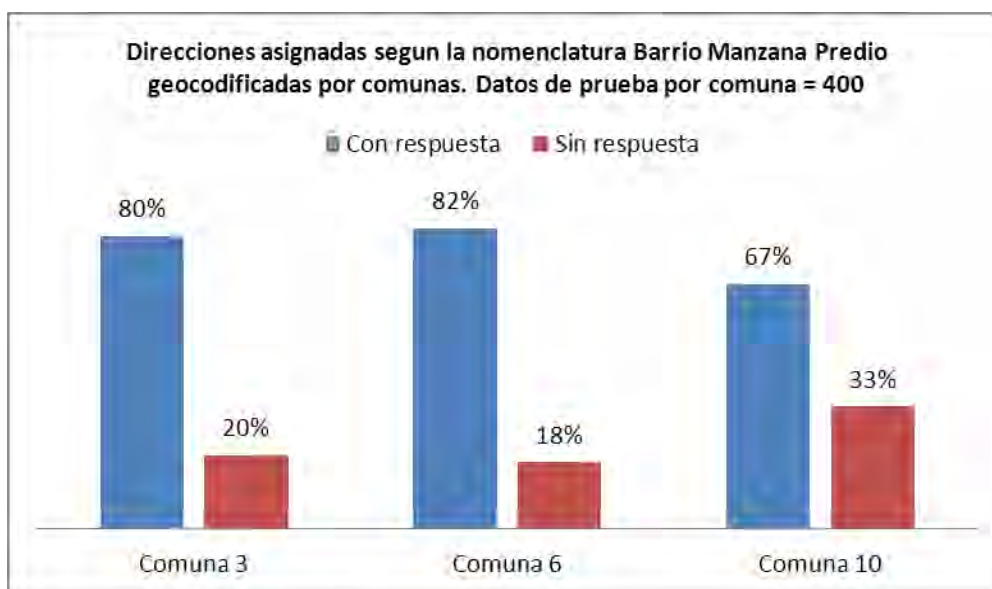


**Caso de prueba 4. Geocodificación de direcciones urbanas ubicadas en la comuna 3. Caso nomenclaturas barrio manzana predio:** El objetivo de esta prueba fue verificar los resultados arrojados por el procedimiento almacenado de geocodificación al ejecutarse con direcciones urbanas registradas en las comunas 3, 6 y 10, comunas con mayor cobertura de direcciones urbanas asignadas según la nomenclatura barrio manzana predio. Se seleccionaron 600 datos de direcciones urbanas con los componentes mínimos de este caso de direcciones urbanas. Los resultados de las pruebas pueden observarse en la tabla 11 y figura 49.

Tabla 11. Resultados de geocodificación de direcciones asignadas según la nomenclatura Barrio Manzana Predio ejecutado por comuna.

	<b>Con respuesta</b>		<b>Sin respuesta</b>	
<b>Comuna 3</b>	479	80%	121	20%
<b>Comuna 6</b>	491	82%	109	18%
<b>Comuna 10</b>	400	67%	200	33%

Figura 49. Resultado geocodificación de direcciones urbanas según la nomenclatura Barrio Manzana Predio



**2.3.5. Identificación y visualización geográfica de patrones delictivos.** Con la disposición del geocodificador de direcciones urbanas para el municipio de Pasto, se planteó la posibilidad de descubrir conocimiento mediante la extracción de patrones o tendencias de los datos registrados por el observatorio del delito del municipio de Pasto considerando la información geográfica agregada por esta herramienta y analizando los patrones obtenidos desde su componente geográfico. A continuación se describen las etapas del procedimiento ejecutado:

**I. Selección y pre procesamiento de datos:** En esta primera etapa, se seleccionaron los datos de las lesiones de causa externa no fatales que ocurrieron entre los años 2002 y 2014, con direcciones asignadas según la malla vial del municipio de Pasto. Como se mencionó

anteriormente, los atributos que corresponden a la información urbana de cada evento delictivo fueron obtenidos con la ejecución del geocodificador de direcciones urbanas. A partir de estos nuevos datos, se descartaron aquellos eventos delictivos que no contenían coordenadas geográficas. La tabla 1 muestra los atributos seleccionados.

Tabla 12. Atributos de los eventos delictivos no fatales seleccionados

N	Atributo	Descripción
1	tipo_delito	Tipo de lesión de causa externa no fatal
2	Día	Día de ocurrencia del delito, calculado a partir de su fecha de ocurrencia
3	trimestre	Trimestre del año en el que ocurrió el delito, calculado a partir de su fecha de ocurrencia
4	barrio	Barrio donde ocurrió el delito, dato asignado por el geocodificador de direcciones.
5	comuna	Comuna donde ocurrió el delito, dato asignado por el geocodificador de direcciones.
6	longitud	Valor numérico decimal, dato asignado por el geocodificador de direcciones.
7	latitud	Valor numérico decimal, dato asignado por el geocodificador de direcciones.

En total se encontraron 8.911 registros que cumplían con los requerimientos de información establecidos. Estos datos fueron recuperados y almacenados en archivos con formato CSV (*comma separated values*) mediante la ejecución de una sentencia escrita en el lenguaje de programación SQL del gestor de bases de datos PostgreSQL.

**II. Extracción y análisis de patrones:** Esta etapa tiene como objetivo descubrir patrones insospechados y de interés en los datos seleccionados y analizar los resultados obtenidos. Para lograrlo, se aplicó la técnica de minería de datos basada en la segmentación o *clustering* de datos [40] [41], lo cual permitió encontrar segmentos o *clusters* con registros que comparten un número de propiedades similares. Fue aplicado el algoritmo de *clustering k-means* implementado en la herramienta de minería de datos Weka [42]. Se configuro el parámetro k del algoritmo *k-means* con el valor 4 con el fin de encontrar 4 grupos o *clusters* que corresponden a los patrones buscados. Los resultados de la ejecución del algoritmo *k-means* que se muestran en la tabla 13, corresponden a los valores de los atributos predominantes de los *clusters* encontrados.

Tabla 13. *Clusters* encontrados con la ejecución del algoritmo k-means

Cluster	Cluster0	Cluster1	Cluster2	Cluster3
Cantidad	2.425	2.870	1.808	1.808
Porcentaje	27	33	20	20

Atributo	Cluster0	Cluster1	Cluster2	Cluster3
tipo_delito	No intencional	No intencional	No intencional	Violencia intrafamiliar
día	Martes	Lunes	Miércoles	Domingo
trimestre	II	III	III	IV
barrio	Aire libre	Centro	El Pilar	El Progreso
comuna	Comuna 2	Comuna 1	Comuna 5	Comuna 5
longitud	-8602800.89	-8602981.9411	-9602257.8476	-8602398.7591
latitud	134700.2291	135385.6017	133695.6267	134188.3962

De acuerdo con la tabla 13, en el cluster0 se identifican las características predominantes del 27% de los delitos seleccionados, las cuales se asocian a delitos de tipo no intencional que ocurren los días martes del segundo trimestre del año, en la comuna 2, específicamente en el barrio Aire Libre. En el cluster1 se identifican las características predominantes del 33% de los delitos procesados, las cuales se asocian a delitos de tipo no intencional que ocurren los días lunes del tercer trimestre del año, en la comuna 1, específicamente en la zona centro de la ciudad. En el cluster2 se identifican las características predominantes del 20% de los delitos seleccionados, las cuales se asocian a delitos de tipo no intencional que ocurren los días miércoles del tercer trimestre del año, en la comuna 5, específicamente en el barrio el Pilar. Finalmente, en el cluster3 se identifican las características predominantes del 20% de los delitos seleccionados, las cuales se asocian a delitos de tipo violencia intrafamiliar, que ocurren los días domingo del cuarto trimestre del año, en la comuna 5, específicamente en el barrio El Progreso.

**III. Visualización geográfica de patrones:** El objetivo de esta etapa fue visualizar e identificar la ubicación geográfica de los *clusters* encontrados a partir de las coordenadas geográficas latitud y longitud de cada registro procesado para realizar su posterior análisis.

Para hacer esto posible, se utilizó la funcionalidad de visualización de asignación de *clusters* implementada en Weka [41], almacenando las instancias o registros visibles en un archivo con formato ARFF (*Attribute Relation File Format*). A su vez, este archivo fue subido de nuevo a Weka para almacenar los datos en un archivo con formato CSV. Como resultado, se obtuvo un archivo

con los registros que fueron seleccionados en la etapa de selección y pre-procesamiento de datos con atributos adicionados por el algoritmo k-means, entre los cuales se encuentra el identificador de instancia o registro visualizado en Weka (*instance\_number*) y el *cluster* al que fue asignado cada registro. Algunos de estos datos pueden observarse en la tabla 14.

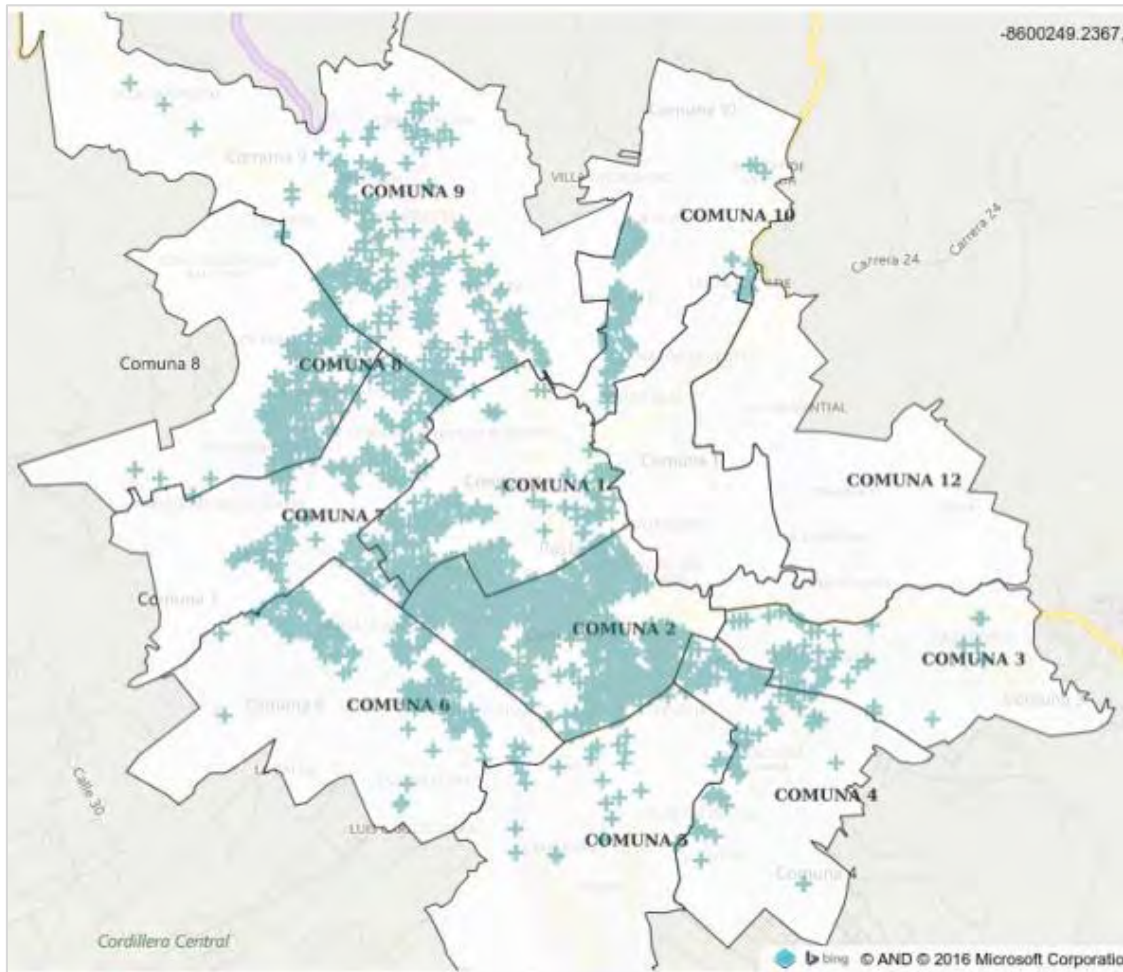
Tabla 14. Algunos de los datos procesados por el algoritmo k-means en Weka

instance_n	delito	dia	trimestre	cuadra	barrio	comuna	longitud	latitud	Cluster
0	'NO INTENCIONAL'	Sabado	IV	'KR 12 CL 15'	'LAS VIOLETAS'	'COMUNA 2'	-8602203.75	133795.59	cluster0
1	'NO INTENCIONAL'	Martes	III	'KR 22 A CL 9'	'LOS ALAMOS'	'COMUNA 2'	-8603199.89	134465.78	cluster0
2	'NO INTENCIONAL'	Domingo	I	'KR 37 CL 18'	PALERMO	'COMUNA 9'	-8603252.09	136458.76	cluster1
3	'VIOLENCIA INTRAFAMILIAR'	Domingo	I	'CL 18 KR 11'	FATIMA	'COMUNA 2'	-8601972.12	133894.59	cluster3
4	'NO INTENCIONAL'	Jueves	II	'KR 33 CL 16 A'	MARIDIAZ	'COMUNA 9'	-8603187.75	135869.95	cluster0
5	'NO INTENCIONAL'	Domingo	III	'CL 18 KR 56'	TOROBAJO	'COMUNA 9'	-8604877.66	137417.67	cluster1
6	'NO INTENCIONAL'	Sabado	IV	'KR 18 CL 20'	'EL PRADO'	'COMUNA 2'	-8602170.74	134509.66	cluster0
7	'VIOLENCIA INTERPERSONAL'	Domingo	I	'KR 13 A CL 20'	'EL RECUERDO'	'COMUNA 2'	-8601841.09	134070.93	cluster3
8	INTENCIONAL AUTOINFLINGIDA	Sabado	II	'CL 5 KR 22 D'	SAN JOSE OBRERO'	'COMUNA 1'	-8603486.53	134499.19	cluster0

Para la visualización geográfica de *clusters*, fue utilizada la funcionalidad de procesamiento y agrupación de datos georreferenciados almacenados en archivos CSV implementada en el visor cartográfico PASTO VIEW. A continuación se muestra y describe la visualización de cada *cluster* encontrado.

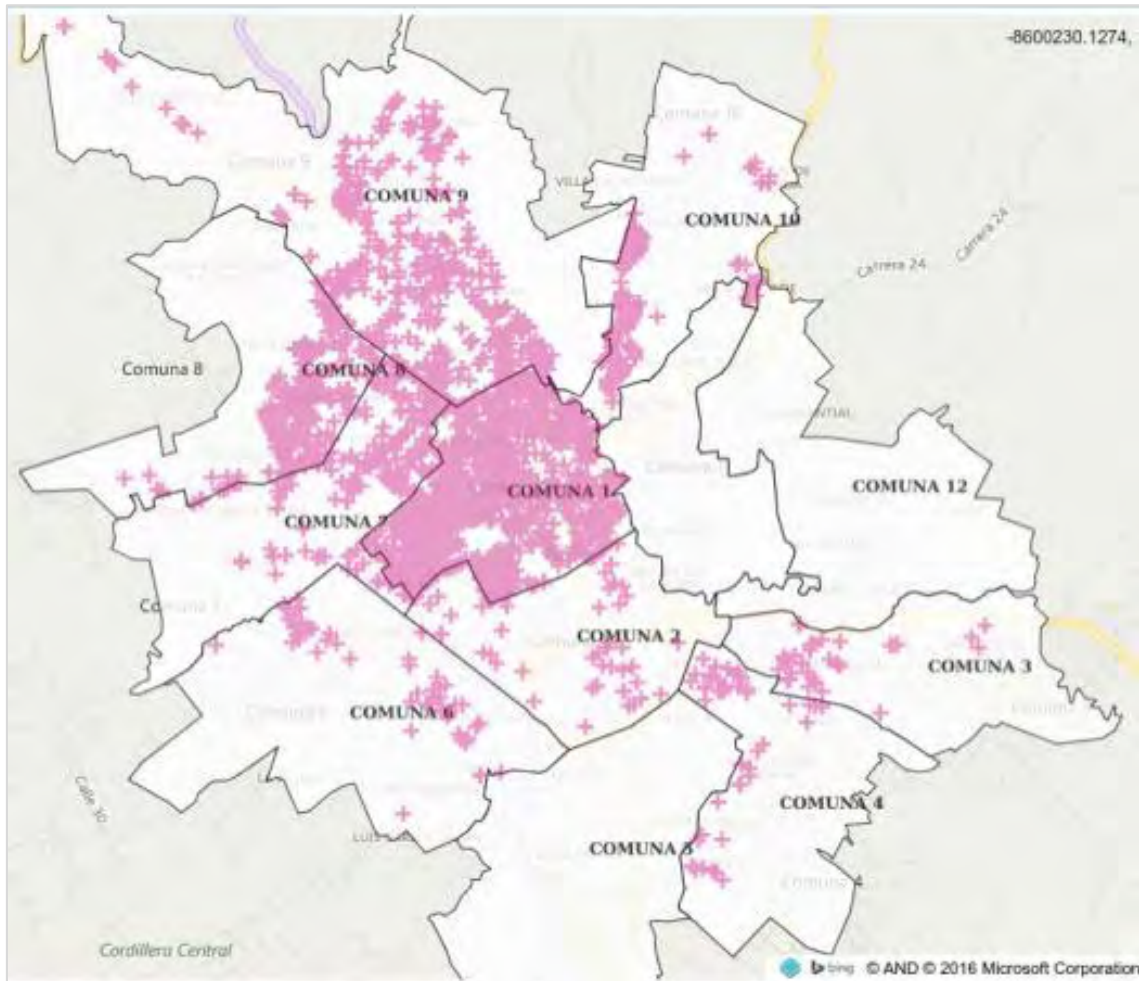
La figura 50 muestra la ubicación geográfica de los eventos delictivos que fueron agrupados en el *cluster0* por el algoritmo *k-means*. Se puede identificar la tendencia descrita por las características predominantes encontradas, ubicando una mayor cantidad de delitos en la comuna 2 del municipio de Pasto. De igual forma se identifican las comunas 8 y 9 con una presencia considerable de delitos.

Figura 50. Visualización geográfica de los eventos delictivos que fueron agrupados al *cluster0* por el algoritmo k-means



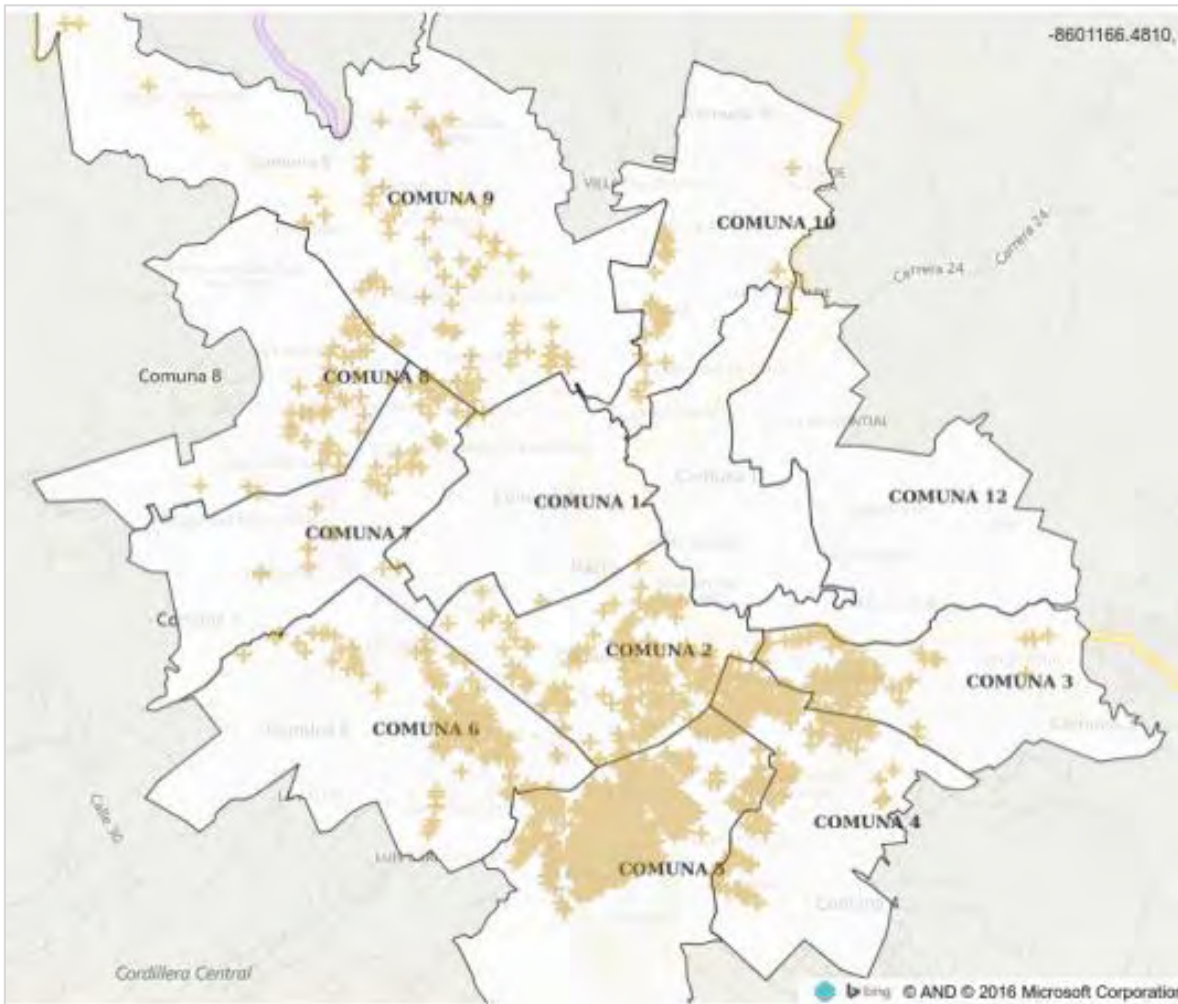
La figura 51 muestra la ubicación geográfica de los eventos delictivos que fueron agrupados en el *cluster1* por el algoritmo *k-means*. Se puede identificar la tendencia descrita por las características predominantes encontradas, ubicando una mayor cantidad de delitos en la comuna 1 del municipio de Pasto. De igual forma se identifican las comunas 7, 8 y 9 con una presencia considerable de delitos.

Figura 51. Visualización geográfica de los eventos delictivos que fueron agrupados al *cluster1* por el algoritmo *k-means*



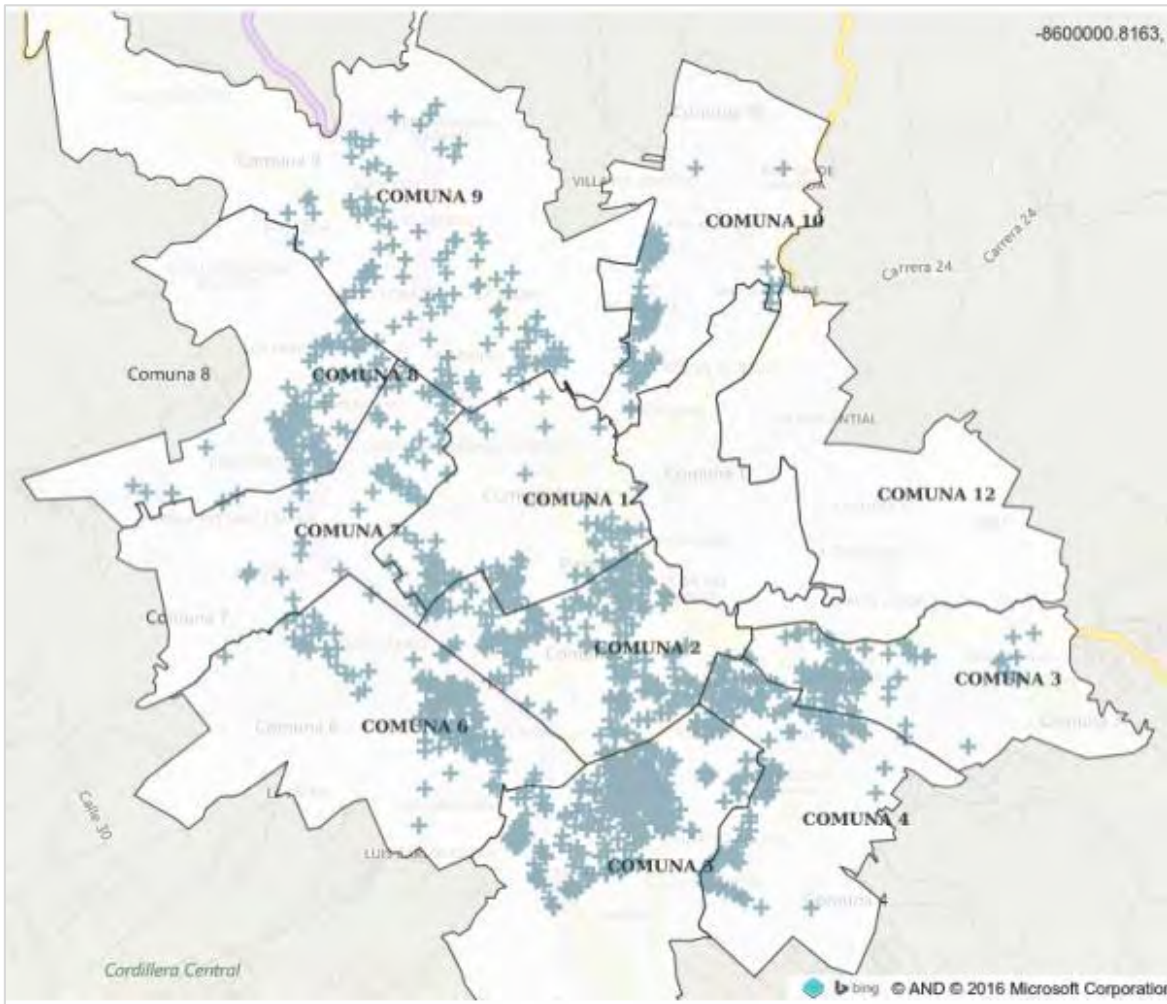
La figura 52 muestra la ubicación geográfica de los eventos delictivos que fueron agrupados en el *cluster2* por el algoritmo *k-means*. Se puede identificar la tendencia descrita por las características predominantes encontradas, ubicando una mayor cantidad de delitos en la comuna 1 del municipio de Pasto. De igual forma se identifican las comunas 1, 2, 3, 4 y 6 con una presencia considerable de delitos.

Figura 52. Visualización geográfica de los eventos delictivos que fueron agrupados al *cluster2* por el algoritmo *k-means*



La figura 53 muestra la ubicación geográfica de los eventos delictivos que fueron agrupados en el *cluster3* por el algoritmo *k-means*. Se puede identificar la tendencia descrita por las características predominantes encontradas, ubicando una mayor cantidad de delitos en la comuna 1 del municipio de Pasto. A diferencia de los anteriores casos, se identifica una alta dispersión de los datos ubicando eventos delictivos en todas las comunas.

Figura 53. Visualización geográfica de los eventos delictivos que fueron agrupados al *cluster3* por el algoritmo *k-means*



En cada caso de *cluster* visualizado geográficamente, se tiene delitos ubicados en una zona distinta a la que describe el patrón, lo cual implica que los registros ubicados en estas zonas poseen características similares, diferenciándose en los atributos geográficos comuna, barrio, longitud y latitud.

### 3. RESULTADOS DE LA INVESTIGACIÓN

La investigación realizada sobre la geocodificación de eventos georreferenciados a nivel de direcciones urbanas en el municipio de Pasto permitió obtener un primer estudio sobre un área de investigación que no ha sido experimentada dentro el contexto del municipio de Pasto.

De igual forma se dio inicio a un área de investigación en el municipio de Pasto que puede ser integrada con otras áreas tales como la georreferenciación de eventos dentro de cualquier contexto de estudio, análisis de datos a nivel de direcciones urbanas, la minería de datos espaciales, entre otros.

A continuación se describen los resultados obtenidos en el desarrollo de esta investigación:

- Repositorio de datos con Información espacial urbana que corresponde a las vías, los barrios y comunas del municipio de Pasto.
- Direcciones urbanas recopiladas en archivos OSM (XML) de las comunas 1, 2, 3, 4, 5 ,6, 7, 8, 9 y 10 del área urbana del municipio de Pasto.
- Geocodificador de direcciones urbanas funcional que permite la búsqueda de la información urbana que corresponde a una dirección ingresada.
- Visor cartográfico acoplado al geocodificador de direcciones, denominado PASTO VIEW que permite la búsqueda y ubicación de direcciones urbanas en el municipio de Pasto y el procesamiento de datos georreferenciados a nivel de direcciones urbanas almacenados en archivos separados por comas CSV.
- Visor cartográfico integrado con el sistema Georreferenciado SIGEODEP, denominado SIGEODEP-SIG que permite visualizar patrones visuales de ocurrencia de delitos a partir de un cruce previo de variables
- Reporte de pruebas realizadas al geocodificador de direcciones urbanas, aplicadas sobre diferentes casos de prueba.
- Reporte de descubrimiento y visualización de patrones delictivos aplicando la técnica *clustering* de minería de datos.

Además, se destaca la publicación del desarrollo de la investigación en los siguientes escenarios:

- Fue presentado un avance de la investigación en el 2do Congreso Andino de Computación, Informática y Educación – CACIED 2015, realizado en la ciudad de Guayaquil – Ecuador. Se obtuvo ideas y sugerencias importantes para el desarrollo de la investigación por parte de docentes investigadores en el área de sistemas de información geográfica asistentes al evento.
- Fue presentado un artículo científico y una ponencia en el V Congreso Internacional de Gestión Tecnológica y de la Innovación – COGESTEC 2016, realizado en la ciudad de Bucaramanga – Colombia, se socializaron aspectos importantes sobre el desarrollo del geocodificador de direcciones urbanas y SIGEODEP SIG con gran acogida por parte de los investigadores asistentes al evento.

#### **4. DISCUSIÓN DE RESULTADOS**

- **Recopilación de direcciones urbanas.**

Dado que la georreferenciación de eventos a nivel de direcciones urbanas requiere de información de referencia con una alta proximidad a la realidad, la ausencia de normas que controlen la asignación de nomenclaturas dificultó la recopilación de información y por ende la geocodificación óptima de estos eventos.

Fueron muchos los barrios en donde se observaron anomalías como cuadras o direcciones repetidas más de una vez, direcciones urbanas con nomenclaturas incorrectas, incorrecta ubicación y especificación de nomenclaturas, incorrecta escritura de las nomenclaturas en cada predio (muchos residentes de la ciudad identifican sus predios de acuerdo con su criterio más no a su asignación real), entre otras. Estas anomalías fueron observadas con una frecuencia alta de ocurrencia en los barrios ubicados en la periferia del municipio de Pasto.

Las direcciones urbanas, cuadras o zonas que presentaron las anomalías descritas fueron descartadas de la recopilación de direcciones para evitar resultados incorrectos en la ejecución del proceso de geocodificación. A pesar de este gran inconveniente se logró recopilar un 70% de las direcciones urbanas del municipio de Pasto.

- **Pruebas de funcionamiento**

Fueron realizados pruebas sobre el geocodificador de direcciones urbanas del municipio de Pasto. Los resultados arrojados en los casos de prueba muestran la efectividad del geocodificador y sus componentes sobre datos de prueba seleccionados de la base de datos del Observatorio del delito del municipio de Pasto.

Las pruebas realizadas sobre el geocodificador de direcciones urbanas asignadas según la malla vial dieron como resultado un rendimiento superior al 50%, el cual se considera satisfactorio en cuanto a la cantidad de direcciones geocodificadas en las comunas 1, 2, 5, 6, 7, 8 y 9 con un porcentaje mayor al 50% de direcciones geocodificadas exitosamente. Las comunas 3, 4 y 10 arrojaron un porcentaje de direcciones geocodificadas menor al 50%, lo cual se relaciona claramente con problemas de asignación de direcciones urbanas presentes en estas comunas.

Las pruebas realizadas al geocodificador de direcciones urbanas asignadas según nomenclatura Barrio – Manzana - Predio arrojaron un rendimiento satisfactorio del proceso de geocodificación ejecutado, logrando georreferenciar más de un 60% de las direcciones de prueba seleccionadas en las comunas como mayor presencia de este tipo de direcciones.

Se destaca la implementación del geocodificador de direcciones urbanas en el lenguaje procedural del sistema gestor de bases de datos PostgreSQL. Esto permitió el procesamiento de grandes lotes de información georreferenciada a nivel de direcciones urbanas, característica que muchos geocodificadores de direcciones urbanas existentes no tienen o no realizan con un rendimiento óptimo.

## 5. CONCLUSIONES

La presente investigación permitió tener un primer acercamiento al área de investigación relacionada con la georreferenciación de eventos dentro del municipio de Pasto. Se planteó una metodología que permitió obtener resultados satisfactorios en cuanto al cumplimiento de los objetivos planteados.

El proyecto de acceso libre a datos espaciales Open Street Maps fue un recurso clave para el desarrollo de la presente investigación. La disposición de los datos de referencia sobre el área urbana del municipio de Pasto almacenados en el servicio de publicación de datos ofrecido por esta organización internacional, la disposición de una herramienta de uso intuitivo para la edición de estos datos como lo fue Java Open Street Editor, en conjunto con las fuentes de información de referencia seleccionadas para la corrección de información urbana y la recopilación de direcciones hicieron posible la construcción del repositorio de información urbanas del municipio de Pasto.

La recopilación de direcciones urbanas utilizando la extensión Address Interpolation permitió encontrar y considerar dificultades como la asignación errónea de direcciones urbanas en gran parte del área urbana del municipio de Pasto, una de las razones por las que el geocodificador de direcciones no tuvo un rendimiento del 100% en las pruebas de funcionamiento realizadas.

Con la implementación de la teoría de autómatas finitos bajo un enfoque determinista para el reconocimiento y limpieza de las direcciones almacenadas en el observatorio del delito se dispuso de un procedimiento almacenado de búsqueda con alta precisión dentro de la información urbana de referencia. Esta implementación posibilitará el uso del geocodificador de direcciones urbanas en posteriores investigaciones relacionadas con la georreferenciación de eventos a nivel de direcciones urbanas.

La captura de datos soportada mediante sistemas de información como SIGEODEP es un problema común que puede alterar notablemente la disposición de datos reales y confiables, la no aplicación de ningún tipo de normas en esta captura da como resultado datos con distintos formatos referenciando no solo direcciones sino también sitios, descripción de sitios, datos adicionales de los eventos como barrio y/o comuna donde ocurrió el evento, descripciones de ausencia de la

dirección de ocurrencia, campos vacíos o nulos entre otros. Por este motivo, fueron descartados una gran cantidad de eventos delictivos por el geocodificador de direcciones urbanas.

Con la investigación realizada sobre el caso de estudio Observatorio del delito del municipio de Pasto se definieron procesos de geocodificación de direcciones urbanas para el área urbana del municipio de Pasto, dando soporte a la geocodificación de direcciones urbanas presentes entre los datos registrados por esta entidad. La investigación realizada abre nuevos espacios para proponer investigaciones en casos de estudio donde se requiera geocodificar y analizar datos a nivel de direcciones urbanas.

El geocodificador de direcciones urbanas para el municipio de Pasto soporta la geocodificación a nivel de direcciones basadas en la nomenclatura barrio – manzanas - predio, un tipo de geocodificación poco tratada en la información de referencia consultada en los supuestos teóricos y antecedentes de esta investigación.

El uso de herramientas basadas en software libre para realizar las tareas de recopilación, manipulación y publicación de la información urbana del municipio de Pasto fue una alternativa de gran utilidad que permitió disponer de funcionalidades orientadas al tratamiento de información espacial. Las ventajas de utilizar este tipo de herramientas se ve reflejado en la construcción exitosa de herramientas innovadoras como el geocodificador de direcciones urbanas y el visor cartográfico para el municipio de Pasto.

La visualización de eventos a nivel de direcciones es un recurso de los Sistemas de Información Geográfica que permite el análisis de datos con una precisión cercana a la realidad. SIGEODEP SIG como herramienta de soporte a la toma de decisiones permite identificar patrones visuales que corresponden a zonas donde ocurren gran cantidad de eventos delictivos, realizando el conteo respectivo de delitos.

La asignación de direcciones urbanas debe realizarse mediante un proceso donde se controlen aspectos como la ubicación urbana del predio, la selección correcta de nomenclaturas considerando las calles donde se encuentra el predio ubicado y asignar los elementos posibles que pueda tener la dirección urbana. La ausencia de algún elemento puede generar errores en la correcta geocodificación de direcciones ubicando sus coordenadas geográficas a una gran distancia de su ubicación real.

Los procesos de geocodificación aplicados al presente caso de estudio muestran un alto porcentaje de direcciones urbanas que no fue posible procesarse, esto debido a errores presentes en la escritura de las direcciones urbanas o a la ausencia de la información de referencia. Estas dos causas pueden presentarse en cualquier investigación relacionada con geocodificación de eventos.

Con la disposición del geocodificador de direcciones urbanas fue aplicada una técnica de minería de datos basada en *clustering* o agrupamiento de datos sobre los eventos delictivos no fatales geocodificados a nivel de direcciones urbanas, con lo cual se pudo observar los resultados arrojados por el algoritmo de *clustering k-means* ejecutado desde su componente geográfico, visualizando la ubicación geográfica de los delitos asignados a cada *cluster* encontrado.

La integración de herramientas utilizadas junto con la metodología aplicada dieron como resultado la construcción del geocodificador de direcciones urbanas y los visores cartográficos PASTO VIEW y SIGEODEP SIG, herramientas innovadoras que podrán ser utilizadas en nuevas investigaciones y servirán de soporte a la toma de decisiones tanto a ciudadanos como a entidades del municipio de Pasto.

## 6. RECOMENDACIONES

A continuación se plantean algunas recomendaciones a partir de desarrollo de la presente investigación:

- Una de las dificultades identificadas en la presente investigación fue la ausencia de formatos estandarizados en el registro de direcciones urbanas, lo cual impidió procesar un gran porcentaje de estas. Se recomienda definir formularios de registro para direcciones urbanas del municipio de Pasto que permita almacenar todos sus posibles componentes, con el fin de obtener mejor resultados en la geocodificación de eventos.
- Entidades gubernamentales como el Instituto Geográfico Agustín Codazzi almacenan los datos espaciales del municipio de Pasto, los cuales de ser posible su acceso permitiría tener una mayor eficiencia en la geocodificación de eventos. Esta alternativa finalmente se descarta debido a las políticas que esta entidad gubernamental establece. Por lo tanto se sugiere fomentar la cooperación entre la academia y las entidades gubernamentales con el objetivo de aunar esfuerzos en el desarrollo de investigaciones cuyos resultados beneficiaran a las entidades involucradas como también a nuestra región.
- Con el fin de abrir espacios para plantear investigaciones orientadas a solventar problemas presentes en el municipio de Pasto haciendo uso del análisis de datos georreferenciados a nivel de direcciones, se recomienda definir métodos de acceso, modificación y publicación libre de los datos del repositorio de información urbana que permita la participación colaborativa de investigadores y usuarios interesados.
- Establecer políticas de control de asignación de nomenclaturas para el área rural en el municipio de Pasto que permitan realizar investigaciones relacionadas con la georreferenciación de eventos registrados en estas zonas.
- Definir nuevos requerimientos al visor cartográfico SIGEODEP – SIG que hagan de esta una herramienta con nuevas funcionalidades dispuestas por los Sistemas de Información Geográfica.
- El geocodificador de direcciones urbanas permite acceder a la información urbana espacial asociada a los eventos delictivos almacenados por el observatorio del delito del municipio de Pasto, por lo anterior se recomienda utilizar esta información georreferenciada a nivel de direcciones urbanas en proyectos de investigación orientados al descubrimiento de conocimiento.

- Utilizar el visor PASTO-VIEW y el geocodificador de direcciones como base para otros proyectos relacionados con eventos georreferenciados en el municipio de Pasto

## 7. BIBLIOGRAFÍA

- [1] O. Huisman y R. A. By, *Principles of Geographic Information Systems: An Introductory Textbook*. International Institute for Geo-Information Science and Earth Observation (ITC), 2009.
- [2] V. Olaya, «Sistemas de Información Geográfica Versión 1.0». 2011.
- [3] R. Timarán, A. Barón, A. Hidalgo, y C. Betancourth, «SIGEODEP: Un primer paso para la Detección de Patrones Delictivos con Técnicas de Minería de Datos.», presentado en IX Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2012), Lima, Perú, 2012.
- [4] Refrations Reseach, «PostGIS Extras, Tiger Geocoder», *Chapter 12. PostGIS Extras*, 2004. [En línea]. Disponible en: <http://postgis.net/docs/Extras.html>. [Accedido: 01-abr-2015].
- [5] US Census Bureau, «TIGER Products». [En línea]. Disponible en: <http://www.census.gov/geo/maps-data/data/tiger.html>. [Accedido: 01-abr-2015].
- [6] Instituto Geográfico Nacional de España, «Cartociudad», *Cartociudad*. [En línea]. Disponible en: <http://www.cartociudad.es/portal/>. [Accedido: 01-abr-2015].
- [7] I. Moya Honduvilla y M. Á. Manso Callejo, «Servicio web de Geocodificación para Cartociudad», en *Actas de I Jornadas Ibéricas de Infra-estructuras de Datos Espaciais, IIIDE 2010*, Lisboa, Portugal, 2010.
- [8] Procálculo Prosis S.A., «Mapas.com.co», *Mapas.co, Cartografía e Imágenes Satelitales*, 2013. [En línea]. Disponible en: <http://www.mapas.com.co>. [Accedido: 01-abr-2015].
- [9] F. Benítez, «Mapas.com.co», *ArcGIS - www.mapas.com.co*, 2012. [En línea]. Disponible en: <http://www.arcgis.com/home/item.html?id=b197b9067c534d9b874778cca16c71a0>.
- [10] J. A. V. Benjumea y D. H. Álvarez, «Proceso de Geocodificación de direcciones en la Ciudad de Medellín: una técnica determinística de georreferenciación de direcciones», *Rev. Ing. USBMed*, vol. 4, n.º 1, pp. 6-21, 2013.
- [11] J. Román y I. Solarte, «Geopasto: Un sistema de información geográfica web orientado al apoyo para la toma de decisiones basados en el plan de ordenamiento territorial del municipio de Pasto», *Línea de Software Y Manejo de Información*, Universidad de Nariño, Pasto, Colombia, 2008.
- [12] C. Arteaga y M. Erazo, «Atlas: Herramienta de cartografía web y geocodificación para el desarrollo de sistemas híbridos en áreas urbanas sobre J2EE y PostgreSQL», *Línea de Software Y Manejo de Información*, Universidad de Nariño, Pasto, Colombia, 2008.

- [13] A. K. W. Yeung y G. B. Hall, *Spatial Database Systems: Design, Implementation and Project Management*. Springer Science & Business Media, 2007.
- [14] P. A. Longley, M. Goodchild, D. J. Maguire, y D. W. Rhind, *Geographic Information Systems and Science*. John Wiley & Sons, 2005.
- [15] Servicio Geológico Colombiano, «Guías de Información geocientífica digital, versión preliminar». 1999.
- [16] K. C. Laudon y J. P. Laudon, *Sistemas de información gerencial: administración de la empresa digital*. Pearson Educación, 2012.
- [17] R. O. Obe y L. S. Hsu, *PostGIS in Action*. Manning Publications Company, 2015.
- [18] A. M. E. Körner y J. I. G. Leiva, *Proyecciones cartográficas: manejo y uso*. Universidad Católica, 1997.
- [19] ESRI, «ArcGIS 9, Geocoding Rule Base Developer Guide». 2003.
- [20] S. Gomatam, R. Carter, M. Ariet, y G. Mitchell, «An empirical comparison of record linkage procedures», *Stat. Med.*, vol. 21, n.º 10, pp. 1485–1496, 2002.
- [21] P. A. Zandbergen, «A comparison of address point, parcel and street geocoding techniques», *Comput. Environ. Urban Syst.*, vol. 32, n.º 3, pp. 214-232, may 2008.
- [22] S. F. M. y M. E. Valdez, *Principios básicos de cartografía y cartografía automatizada*. UAEM, 2003.
- [23] J.-M. Kraak y A. Brown, *Web Cartography*. CRC Press, 2003.
- [24] J. Bennett, *OpenStreetMap - Be your own Cartographer*. Birmingham, U.K.: Packt Publishing, 2010.
- [25] Instituto Geográfico Agustín Codazzi, «Propuesta de Estándar de las Direcciones Urbanas para los Equipamientos del Ministerio de Educación.» 2009.
- [26] Instituto Geográfico Agustín Codazzi, «Manual de reconocimiento predial.» 1997.
- [27] Webmaster, «Observatorio del Delito». [En línea]. Disponible en: <http://www.pasto.gov.co/index.php/nuestras-dependencias/secretaria-de-gobierno/observatorio-del-delito>. [Accedido: 27-may-2016].
- [28] «OpenStreetMap», *OpenStreetMap*. [En línea]. Disponible en: <https://www.openstreetmap.org/>. [Accedido: 18-may-2016].
- [29] «Java Open Street Map Editor (JOSM). Editor libre y de código abierto para Open Street Maps.» [En línea]. Disponible en: <https://josm.openstreetmap.de/>. [Accedido: 18-may-2016].
- [30] «Projections | Information | OpenStreetMapData». [En línea]. Disponible en: <http://openstreetmapdata.com/info/projections>. [Accedido: 25-jul-2016].

- [31] «Google Maps», *Google Maps*. [En línea]. Disponible en: <https://www.google.com.co/maps>. [Accedido: 18-may-2016].
- [32] «Instituto Geografico Agustin Codazzi - Visor catastro nacional». [En línea]. Disponible en: <http://geoportal.igac.gov.co/ssigl2.0/visor/galeria.req?mapaId=23>. [Accedido: 18-may-2016].
- [33] «Quantum GIS. Sistema de información Geografica libre y de código abierto.» [En línea]. Disponible en: <http://www.qgis.org/es/site/>. [Accedido: 18-may-2016].
- [34] «JOSM/Plugins/AddrInterpolation - OpenStreetMap Wiki». [En línea]. Disponible en: <http://wiki.openstreetmap.org/wiki/JOSM/Plugins/AddrInterpolation>. [Accedido: 18-may-2016].
- [35] «PostgreSQL: The world's most advanced open source database». [En línea]. Disponible en: <http://www.postgresql.org/>. [Accedido: 18-may-2016].
- [36] «PostGIS — Spatial and Geographic Objects for PostgreSQL». [En línea]. Disponible en: <http://postgis.net/>. [Accedido: 18-may-2016].
- [37] «EPSG:3857 -- WGS84 Web Mercator (Auxiliary Sphere): SR-ORG Projection -- Spatial Reference». [En línea]. Disponible en: <http://spatialreference.org/ref/sr-org/7483/>. [Accedido: 25-jul-2016].
- [38] E. Taveira, «pg\_similarity», *pg\_similarity Documentation*. [En línea]. Disponible en: <http://pgsimilarity.projects.pgfoundry.org/>. [Accedido: 19-sep-2016].
- [39] «GeoServer». [En línea]. Disponible en: <http://geoserver.org/>. [Accedido: 28-may-2016].
- [40] J. H. Orallo, M. J. R. Quintana, y C. F. Ramírez, *Introducción a la minería de datos*. Editorial Alhambra S. A. (SP), 2004.
- [41] M. Abernethy, «Classification and clustering», 11-may-2010. [En línea]. Disponible en: <https://www.ibm.com/developerworks/library/os-weka2/>. [Accedido: 29-sep-2016].
- [42] University of Waikato, «Weka 3 - Data Mining with Open Source Machine Learning Software in Java». [En línea]. Disponible en: <http://www.cs.waikato.ac.nz/ml/weka/>. [Accedido: 27-sep-2016].

# **ANEXOS**

## ANEXO A.

### Manual de digitalización y corrección de direcciones urbanas

A continuación se describe el procedimiento necesario para modificar los archivos OSM que contienen las direcciones urbanas del municipio de Pasto. Los siguientes pasos fueron ejecutados dentro del sistema operativo Linux Ubuntu en su versión 14.04.

#### 1. Instalación editor JOSM desde repositorio de paquetes

El procedimiento de la instalación del editor JOSM puede ser consultado con detalle en el siguiente enlace: <https://josm.openstreetmap.de/wiki/Download>

Inicialmente se debe agregar una línea a la lista de recursos de paquetes ubicada en la ruta `/etc/apt/sources.list` con las siguientes instrucciones:

```
$ sudo nano /etc/apt/source.list  
  
#agregar esta linea al final del archivo  
deb https://josm.openstreetmap.de/apt trusty universe
```

Se guardan los cambios realizados, posteriormente se actualiza la lista de paquetes y se instala el editor JOSM con las siguientes instrucciones:

```
$ sudo apt-get update  
$ sudo apt-get install josm-latest
```

Una vez terminado el proceso de instalación podremos ingresar al aplicativo instalado como se muestra en la figura A1.

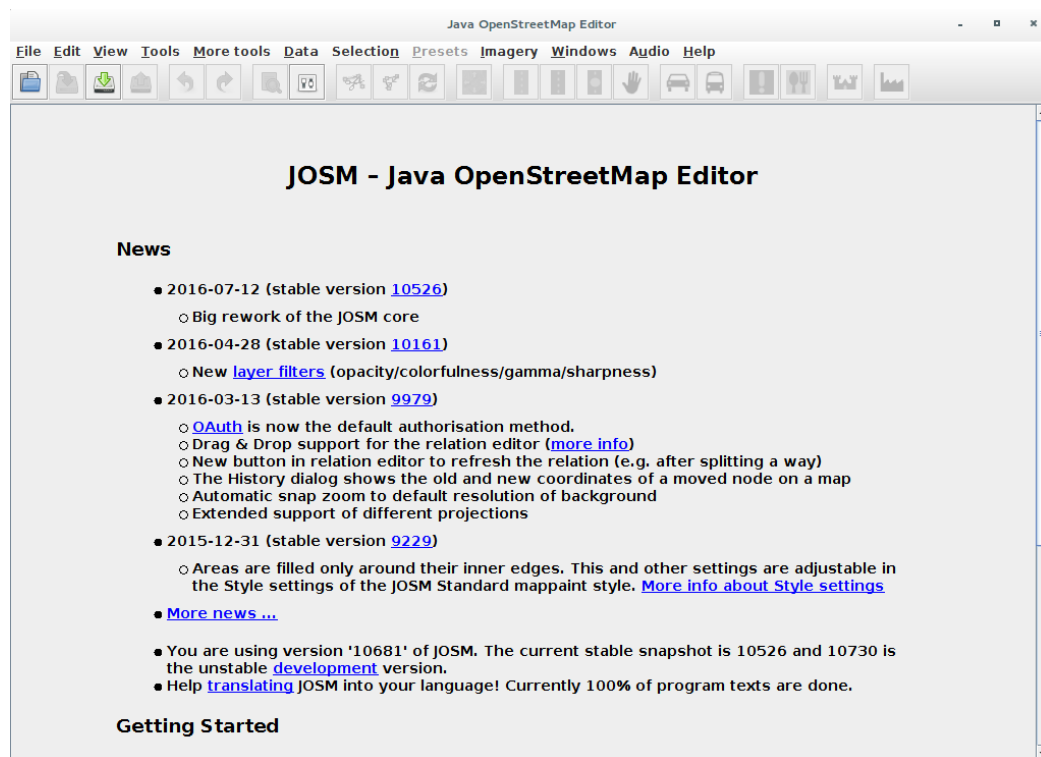


Figura A1. JOSM en ejecución.

Posteriormente instalamos el complemento AddressInterpolation, necesario para la digitalización de direcciones urbanas. En JOSM nos dirigimos al menú Editar/Preferencias/Complementos y buscamos el complemento que deseamos instalar, a continuación seleccionamos la opción Descargar Lista, como se muestra en la figura A2.

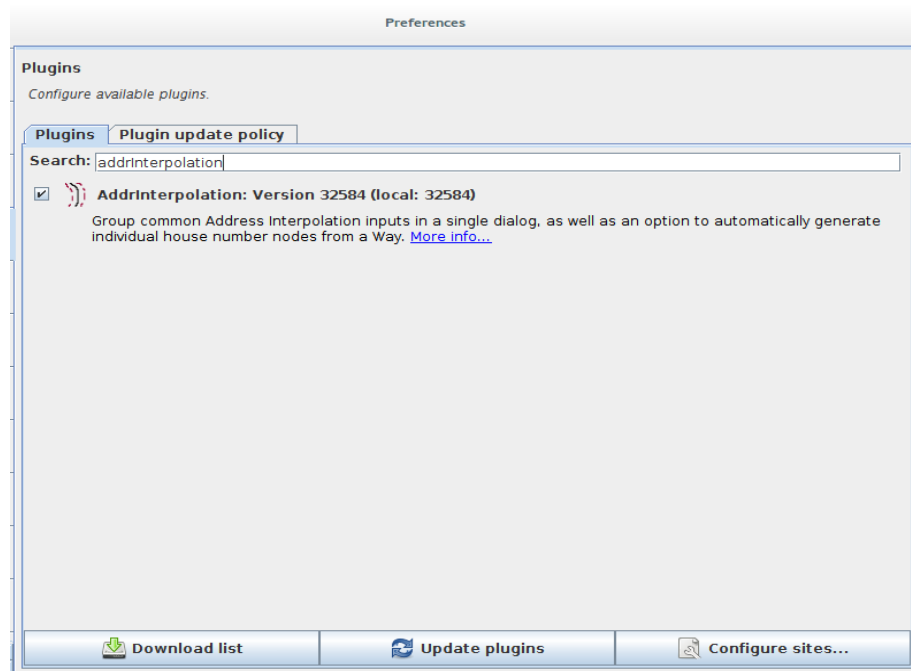


Figura A2. Instalación complemento AddressInterpolation en JOSM

Una vez instaladas las herramientas necesarias para digitalizar direcciones abrimos el archivo OSM que deseamos editar o corregir. Como ejemplo en este tutorial se utilizará el archivo OSM con las direcciones urbanas de la comuna 10. En el menú Archivo/Abrir buscamos y seleccionamos el archivo deseado, el cual se cargará en el aplicativo como se muestra en la figura A3.

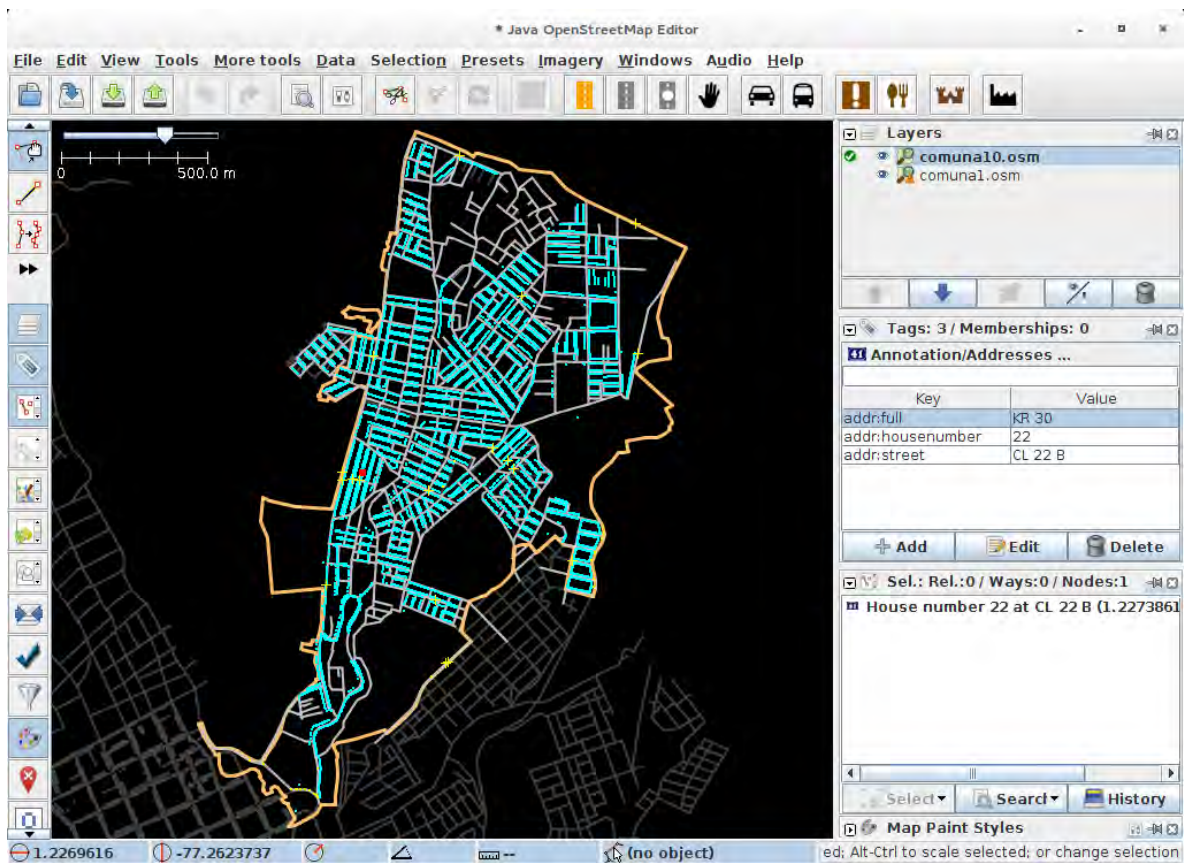


Figura A3. Archivo OSM con direcciones urbanas de la Comuna 10 cargados en JOSM

El archivo OSM seleccionado se cargará como una capa de datos espaciales, en este caso de geometrías basadas en puntos. Al seleccionar cualquiera de estos puntos se puede acceder a funcionalidades como moverlos en el espacio y visualizar sus características. Las características de cada punto cargado pueden verse en el panel derecho de la aplicación habilitando el subpanel Etiquetas (Tags). Estos valores pueden editarse dando doble clic sobre la Etiqueta que desea modificarse, como se muestra en la figura A4.

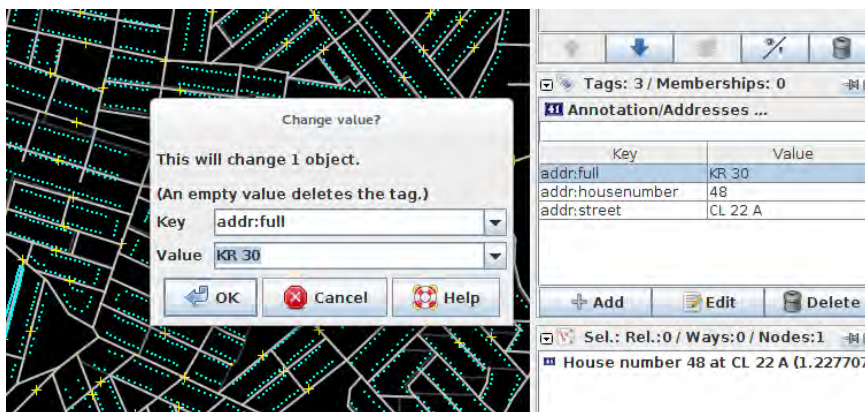


Figura A4. Edición de los atributos de una geometría (Punto).

## 2. Atributos de direcciones urbanas en JOSM

Antes de digitalizar o corregir los archivos OSM con direcciones urbanas del municipio de Pasto se debe considerar los atributos o componentes que deben ingresarse. Dado que el complemento AddressInterpolation esta creado para la digitalización de direcciones de Estados Unidos, para el contexto del municipio de Pasto se hizo una adaptación como se describe a continuación:

Para el caso de direcciones asignadas según la malla vial cada punto debe tener las siguientes etiquetas (véase figura A5):

- *addr:street* es la etiqueta que almacena la vía principal de la dirección urbana, esta debe iniciar con la nomenclatura estandarizada correspondiente: CL (Calle), KR (Carrera) ò DG (Diagonal). Las letras y/o sufijos que acompañen al identificador de calle deben escribirse con espacio y de acuerdo al orden establecido por la propuesta de estandarización de direcciones urbanas. Ejemplos: KR 3, KR 3 A, KR 3 A BIS, KR 3 A SUR.
- *addr:full* es la etiqueta que almacena la vía generadora o secundaria de la dirección urbana, esta debe iniciar con la nomenclatura estandarizada correspondiente: CL (Calle), KR (Carrera) ò DG (Diagonal). Las letras y/o sufijos que acompañen al identificador de calle deben escribirse con espacio y de acuerdo al orden establecido por la propuesta de estandarización de direcciones urbanas. Ejemplos: CL 20, CL 20 A, CL 20 A BIS, CL 20 A SUR.
- *addr:housenumber* es la etiqueta que almacena la distancia de la dirección urbana. Este campo solo debe contener valores numéricos.

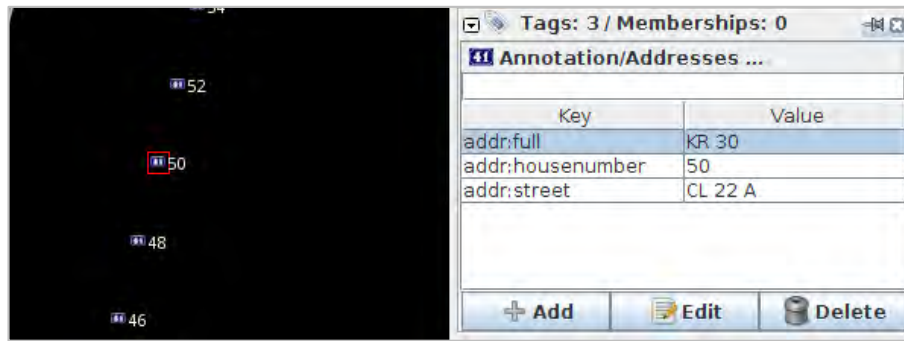


Figura A5. Etiquetas de un punto que corresponde a una dirección asignada según la malla vial

Para el caso de direcciones asignadas según la nomenclatura barrio – manzana – predio cada punto debe tener las siguientes etiquetas (véase figura A6):

- *addr:full* es la etiqueta que almacena el identificador de la manzana (block), este debe siempre iniciar con la nomenclatura MZ En caso de que este identificador contenga números y letras, este último elemento debe escribirse en mayúscula y sin espacio. Ejemplos: MZ 10, MZ 10A, MZ F, MZ F10,
- *addr:housenumber* es la etiqueta que almacena el identificador de predio (houseid), este debe solo contener este valor sin ninguna nomenclatura adicional. En caso de que este identificador contenga números y letras, este último elemento debe escribirse en mayúscula y sin espacio. Ejemplos: 11, 11A.

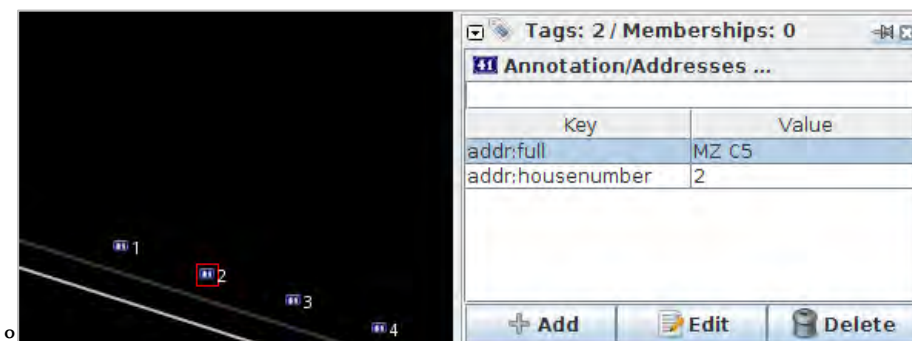


Figura A6. Etiquetas de un punto que corresponde a una dirección asignada según la nomenclatura Barrio Manzana Predio

### 3. Carga de capas base

Tanto para la corrección (modificación) como para la digitalización de nuevas direcciones es necesario tener como referencia una capa base que permita visualizar el mapa del área urbana del municipio de Pasto. En el menú Imágenes (véase figura A7) seleccionamos Bing Maps, mapa que fue utilizado para digitalizar la mayor parte de direcciones urbanas.

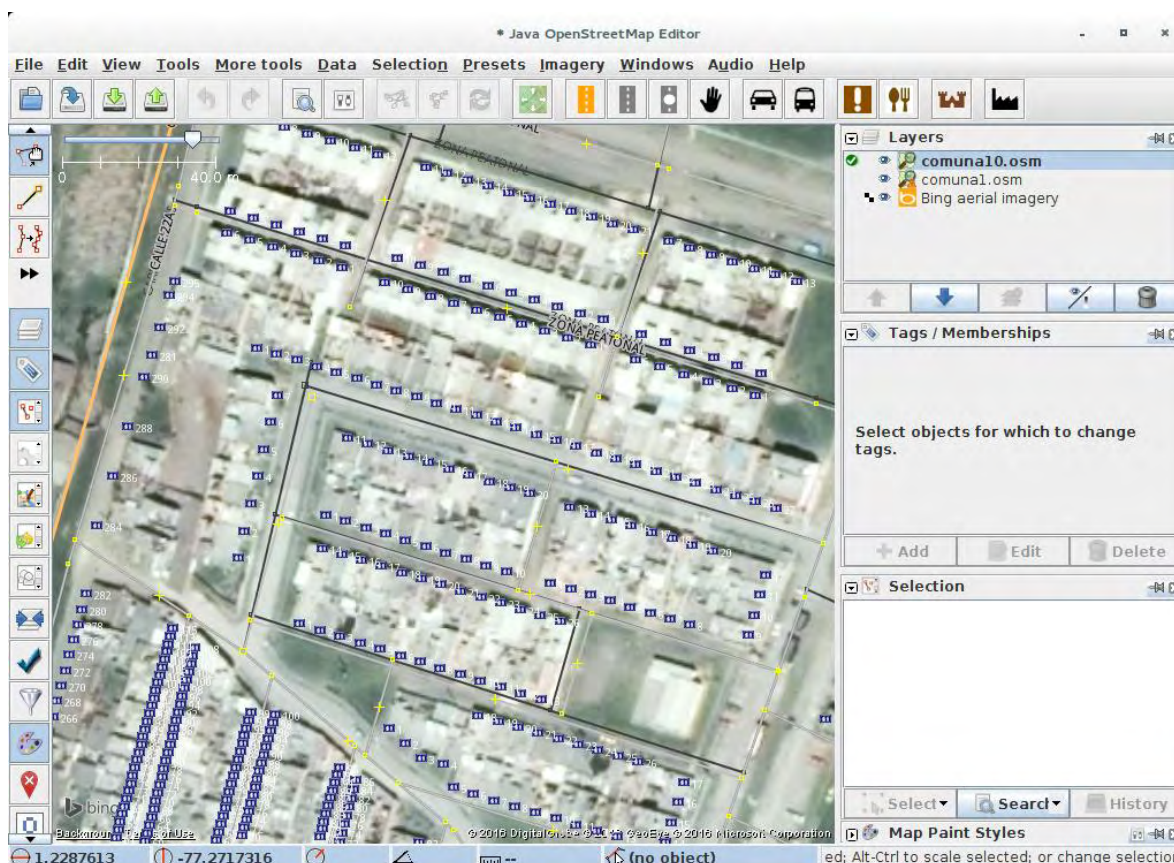


Figura A7. Bing Maps cargado como capa base en JOSM

### 4. Digitalización de direcciones urbanas

Para la digitalización de direcciones urbanas utilizaremos el complemento AddressInterpolation previamente instalado. Este permite acceder a una funcionalidad muy útil para la recolección de direcciones urbanas, como lo es la interpolación de puntos a partir de una geometría o línea base. Como resultado se tiene un conjunto de puntos generados a partir de una línea dibujada, cada punto tendrá los atributos de las direcciones urbanas según se especifique en la previa configuración de la interpolación.

Inicialmente identificamos una cuadra donde deseamos crear un conjunto de puntos que representarán direcciones urbanas ubicadas en la zona seleccionada. Con ayuda de la malla vial dibujamos una línea paralela a la vía principal que atraviesa la cuadra seleccionada. Manteniendo seleccionada la tecla Ctrl, seleccionamos tanto la vía principal como la línea paralela que será interpolada (véase figura A8).

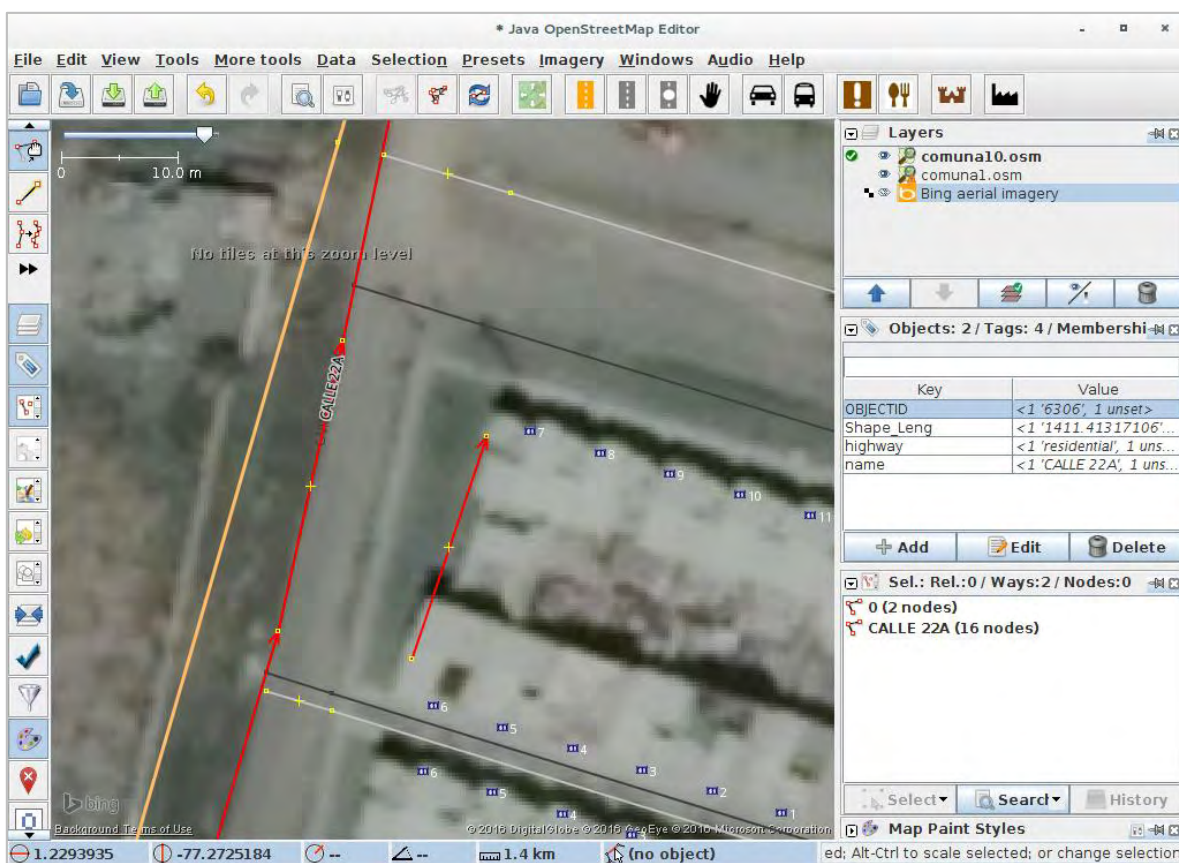


Figura A8. Selección de vía principal (vía de mayor longitud) y línea dibujada, ambas de color rojo.

Una vez seleccionadas estas geometrías, en el menú Datos seleccionamos la opción AddressInterpolation, la cual nos dirige al cuadro de configuración de complemento (véase figura A9). Para mayor comodidad se recomienda acceder desde la combinación de teclas Ctrl + Alt + Z.

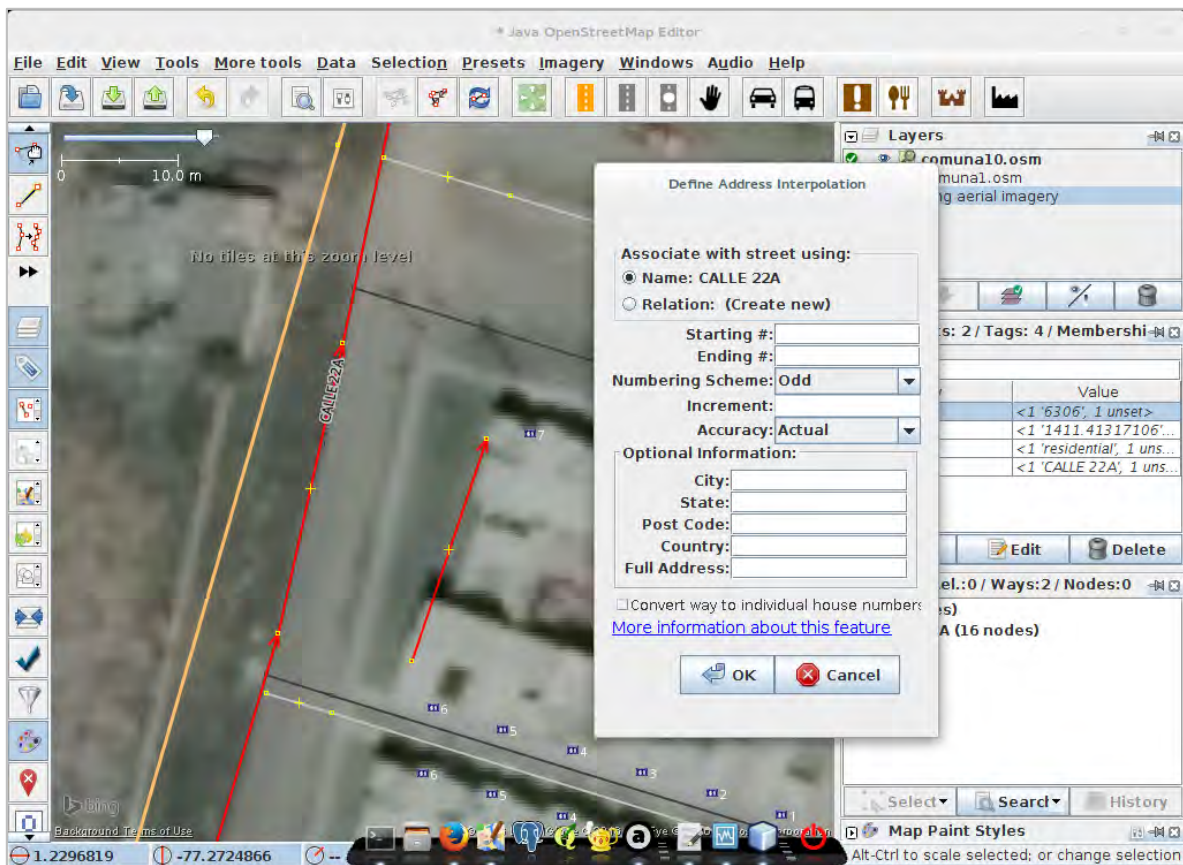


Figura A9. Complemento AddressInterpolation en ejecución

A continuación se describen las principales opciones de este complemento que deben configurarse para obtener correctamente un conjunto de puntos que representan las direcciones urbanas del municipio de Pasto.

Para el caso de direcciones urbanas asignadas según la malla vial se tiene la siguiente configuración:

- La opción *Name* corresponde a la etiqueta `addr:street`
- Las opciones *Inicio (Starting)* y *Fin (Ending)* corresponden al componente distancia de las direcciones. Cabe aclarar que estos valores deben configurarse según la definición de direcciones de la cuadra a digitalizar. De acuerdo a la norma de asignación de direcciones del Instituto Geográfico Agustín Codazzi, un lado de la cuadra tendrá valores pares y el otro tendrá valores impares.

- El esquema de numeración define si la numeración de los puntos generados con la interpolación serán pares o impares.
- La opción *Full Address* corresponde al campo *addr:full*
- Debe habilitarse la opción *Convertir vía en direcciones individuales*.
- Los demás campos deben dejarse vacíos.

La figura A10 muestra un ejemplo de configuración de estos valores en el cuadro de configuración de AddressInterpolation. La figura A11 muestra el resultado de la interpolación de puntos generando direcciones urbanas.

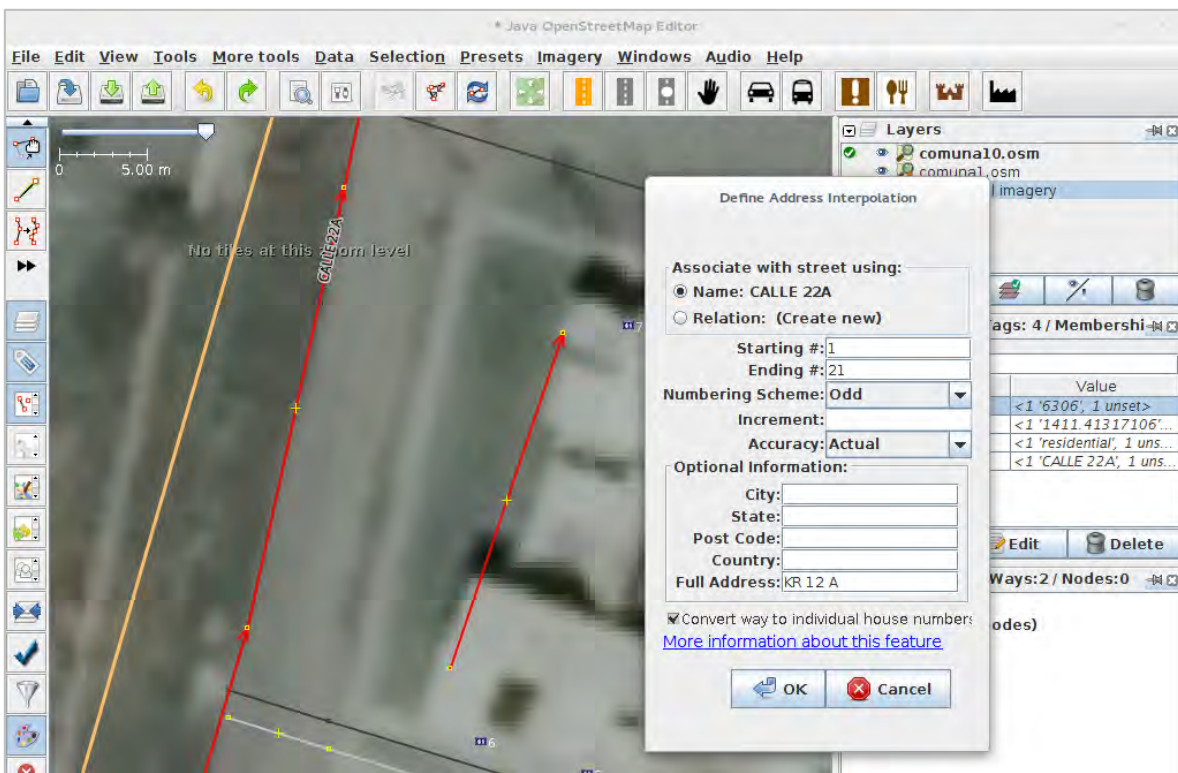


Figura A10. Configuración de valores para generar puntos iniciando desde la distancia 1 y terminando en la distancia 21 (Impar).

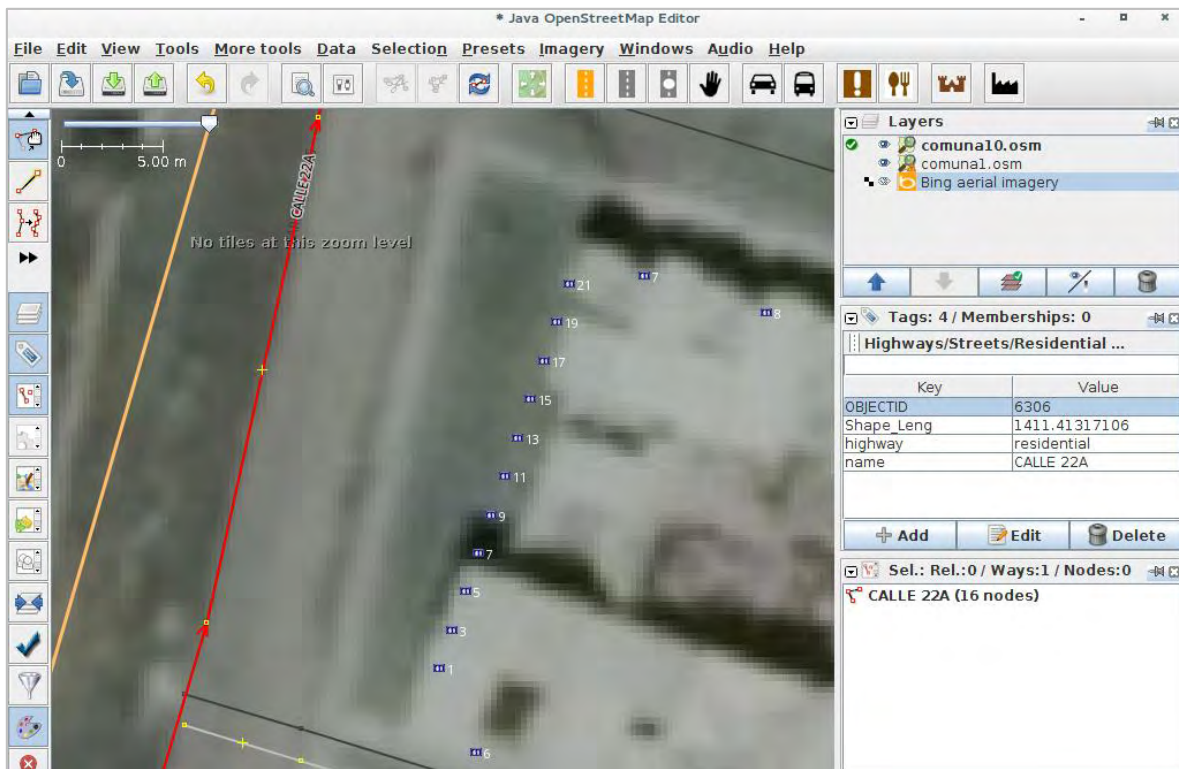


Figura A11. Puntos generados con el complemento AddressInterpolation

Para el caso de direcciones urbanas asignadas según la nomenclatura Barrio Manzana Predio se tiene la siguiente configuración:

- La opción *Name* corresponde a la etiqueta *addr:street*, para este caso de direcciones urbanas debe eliminarse de los puntos generados
- Las opciones *Inicio (Starting)* y *Fin (Ending)* corresponden al componente distancia de las direcciones. Cabe aclarar que estos valores deben configurarse según la definición de direcciones de la cuadra a digitalizar.
- El esquema de numeración debe definirse en la opción Todos, con lo cual se definirán valores continuos con incremento en 1
- La opción *Full Address* corresponde al campo *addr:full*
- Debe habilitarse la opción Convertir vía en direcciones individuales.
- Los demás campos deben dejarse vacíos.

La figura A12 muestra un ejemplo de configuración de estos valores en el cuadro de configuración de AddressInterpolation. La figura A13 muestra el resultado de la interpolación de puntos generando para este caso direcciones urbanas.

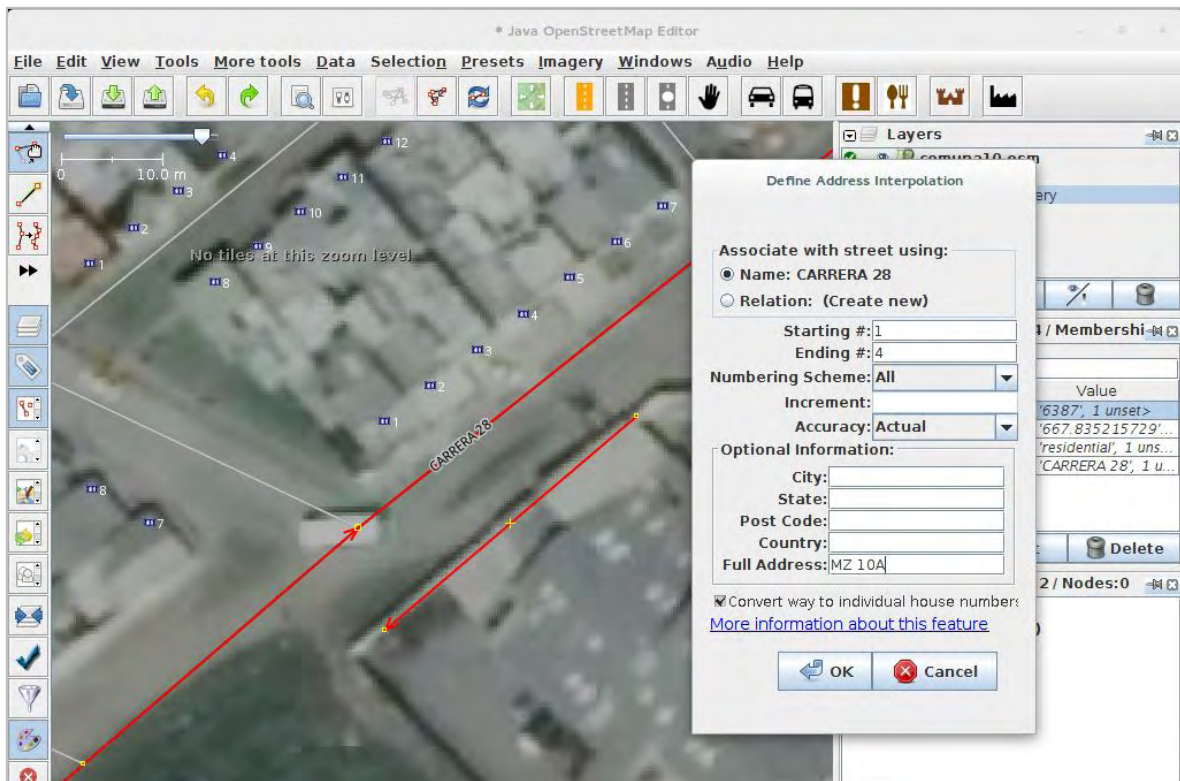


Figura A12. Configuración de valores para generar puntos iniciando desde el predio con identificación 1 hasta el predio con identificación 4.

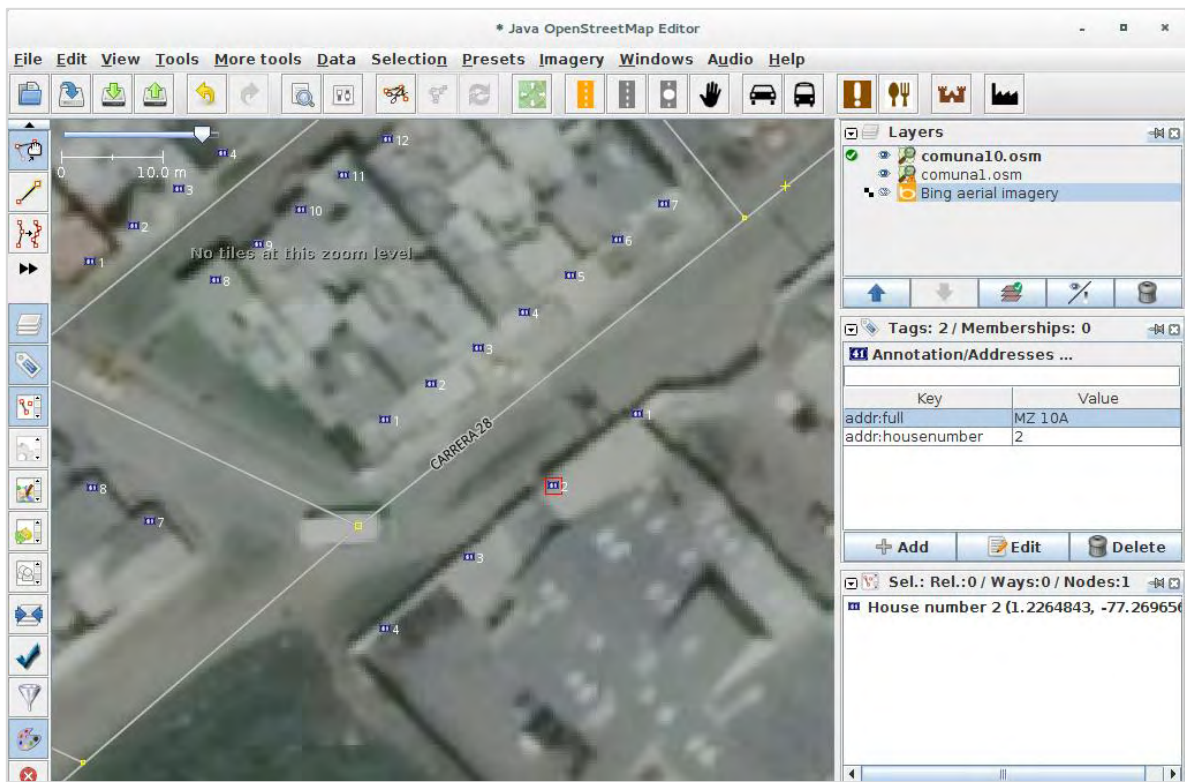


Figura A13. Puntos generados y sus etiquetas.

## ANEXO B.

### Instalación del geocodificador de direcciones urbanas para el municipio de Pasto.

A continuación se describen los requerimientos mínimos, instalación y uso del geocodificador de direcciones urbanas. Los pasos descritos fueron ejecutados sobre la plataforma de sistema operativo Linux/Ubuntu en su versión 14.04. Los archivos requeridos para la instalación pueden descargarse del repositorio <http://grias.udenar.edu.co/owncloud>

#### 1. Software requerido:

##### 1.1. Gestor de bases de datos PostgreSQL con extensión PostGIS

El geocodificador fue implementado con el gestor PostgreSQL en su versión 9.4. Para el manejo de datos espaciales fue utilizado PostGIS en su versión 2.1. El acceso y manipulación de la base de datos fue realizada con el entorno Pgadmin3.

Instalación librerías requeridas:

```
$ sudo apt-get install libxml2 libxml2-dev libproj-dev libgeos-3.4.2  
libgdal-dev libjson0 libjson0-dev
```

Instalación del gestor PostgreSQL 9.4 - PostGIS 2.1 - Pgadmin3:

```
$ sudo apt-get install postgresql-9.4 postgresql-9.4-postgis-2.1  
postgresql-client-9.4 postgresql-server-dev-9.4 postgresql-contrib-9.4  
pgadmin3
```

Configuración de la contraseña del usuario postgres:

```
$ su postgres  
$ psql postgres
```

```
=# ALTER USER 'postgres' WITH PASSWORD 'contraseña'  
=# \q
```

## 1.2. Extensión para PostgreSQL pg\_similarity

Fue utilizado el algoritmo jarowinkler de la extensión pg\_similarity, permitiendo calcular la similitud entre dos cadenas de texto. Esta funcionalidad es utilizada en la estandarización del barrio ingresado junto con la dirección asignada según la nomenclatura barrio – manzana - predio.

Instalación:

```
$ cd pg_similarity  
$ USE_PGXS=1 make  
$ USE_PGXS=1 make install
```

## 1.3. Librerías de Python para la migración de datos OSM

Para la migración de datos OSM a PostgreSQL se utilizó la plataforma del lenguaje de programación Python en su versión 2.7, el adaptador de bases de datos Psycopg2 para la conexión con PostgreSQL y la librería GDAL (Geospatial Data Abstraction Library) para el manejo de las geometrías.

Instalación Python 2.7 – Psycopg2 – GDAL:

```
$ sudo apt-get install python2.7 python-psycopg2 python-gdal
```

## 2. Instalación del esquema de datos y funciones de geocodificación de direcciones:

Cabe mencionar que esta instalación puede realizarse en una base de datos nueva o existente, ya que el geocodificador se instalará en un esquema de datos independiente llamado geocoder. Para instalar este esquema de datos y las funciones del geocodificador debe ejecutarse las siguientes líneas de comandos:

```
$ su postgres  
$ psql -c 'CREATE DATABASE geocoder;'  
$ psql -d geocoder -c '\i  
/<ruta_archivos>/geocoder/install_geocoder.sql'
```

### 3. Migración de datos OSM a PostgreSQL

Para la migración de los datos almacenados en los archivos OSM se requieren de los scripts codificados en los archivos **polygon\_linestring\_migration.py** y **points\_migration.py**

**3.1. Migración de barrios y comunas:** El script **polygon\_linestring\_migration.py** se encarga de leer los archivos OSM que corresponden a los barrios y comunas del municipio de Pasto, recorrer la estructura de datos XML almacenada, identificar los datos de cada barrio o comuna e insertarlos como nuevos registro en la relación barrios o comunas según se haya especificado. A continuación se especifica la línea de ejecución por terminal con sus posibles parámetros:

```
$ ./<ruta_archivo>/polygon_linestring_migration.py <parámetros>
```

Donde los parámetros ingresados pueden ser:

- -f <ruta del archivo>: Ruta donde se encuentra el archivo OSM, valor obligatorio.
- -h <dirección IP>: Dirección IP del servidor donde está instalado el gestor PostgreSQL. Por defecto está configurado localhost.
- -p <puerto>: Puerto de conexión con el gestor PostgreSQL. Por defecto está configurado el puerto 5432
- -d <nombre base de datos>: Nombre de la base de datos donde está instalado el esquema de datos. Este valor es obligatorio.
- -u <nombre usuario>: Nombre del usuario que administra la base de datos. Por defecto está configurado postgres.
- -w <contraseña>: Contraseña del usuario ingresado, valor obligatorio.
- -o <id\_tipo\_geometría>: Selección de un tipo de geometría. Los valores permitidos son l (geometría tipo Linestring) o p (geometría tipo Polygon).
- -t <id\_tabla\_destino>: Selección de la tabla donde se almacenará cada geometría procesada con sus atributos. Los valores permitidos son: c (tabla comunas), n (tabla barrios), r (tabla vías).

La instrucción de ejecución para migrar los datos que corresponden a las comunas es la siguiente:

```
$ ./<ruta_archivo>/polygon_linestring_migration.py -f data/communes.osm  
-d geocoder -u postgres -w 1234 -o p -t c
```

La instrucción de ejecución para migrar los datos que corresponden a los barrios es la siguiente:

```
$ ./<ruta_archivo>/polygon_linestring_migration.py -f  
data/neighborhoods.osm -d geocoder -u postgres -w 1234 -o p -t c
```

La instrucción de ejecución para migrar los datos que corresponden a las vías es la siguiente:

```
$ ./<ruta_archivo>/polygon_linestring_migration.py -f data/roads.osm -d  
geocoder -u postgres -w 1234 -o l -t r
```

**3.2. Migración de direcciones urbanas:** El script `points_migration.py` se encarga de leer los archivos OSM que corresponden a las direcciones urbanas, recorrer la estructura de datos XML almacenada, identificar las direcciones recopiladas según la malla vial y según la nomenclatura barrio - manzana – identificación de predio e insertarlas en su respectiva tabla.

La línea de ejecución por terminal con sus posibles parámetros es la siguiente:

```
$ ./<ruta_archivo>/points_migration.py <parámetros>
```

Donde los parámetros ingresados pueden ser:

- `-f <ruta del archivo>`: Ruta donde se encuentra el archivo OSM, valor obligatorio.
- `-h <dirección IP>`: Dirección IP del servidor donde está instalado el gestor PostgreSQL. Por defecto está configurado localhost.
- `-p <puerto>`: Puerto de conexión con el gestor PostgreSQL. Por defecto está configurado el puerto 5432
- `-d <nombre base de datos>`: Nombre de la base de datos donde está instalado el esquema de datos. Este valor es obligatorio.
- `-u <nombre usuario>`: Nombre del usuario que administra la base de datos. Por defecto está configurado postgres.
- `-w <contraseña>`: Contraseña del usuario ingresado, valor obligatorio.

La instrucción de ejecución para migrar los datos que corresponden a las comunas es la siguiente:

```
$ ./<ruta_archivo>/points_migration.py -f data/addresses.osm -d
geocoder -u geocoder -u postgres -w 1234
```

**4. Asignación de identificadores de barrios y comunas a direcciones urbanas:** Dado que la información sobre el barrio y comuna al que pertenece cada dirección urbana no fue recopilada, esta información se sentencias creadas con el SQL utilizando algunas de las funciones disponibles en la extensión PostGIS para el análisis de relaciones espaciales entre geometrías también llamadas Topologías.

Para actualizar estos datos debe ejecutarse la siguiente instrucción:

```
$ su postgres
$ psql -d geocoder -c "\i /<ruta_archivos>/geocoder/update_fields.sql"
```

**5. Prueba de funcionamiento del geocodificador:**

Para verificar el funcionamiento del geocodificador inicialmente debe ingresarse a la base de datos donde está instalado el geocodificador:

```
$ su postgres
$ psql geocoder;
```

Posteriormente se ejecuta las siguientes consultas SQL, como resultado se tendrá la información urbana correspondiente a la dirección ingresada:

```
=# SELECT geocode_address('kra 20 13 A 45');
                                     geocode_address
-----
 ("KR 20 13 A 45","LAS AMERICAS","COMUNA 1",-602713.06642356,134491.159088312)
```

Si la dirección es correctamente normalizada y encontrada dentro del repositorio de información urbana, la función retornará la información geográfica asociada a esa dirección.

Para acceder a los elementos de forma independiente, se debe ejecutar la siguiente instrucción SQL:

```

=# WITH geocode AS (
    SELECT geocode_address('Kra 20 13 A 45') AS result
)
SELECT
    (result).address AS direccion_normalizada,
    (result).neighborhood AS barrio,
    (result).commune AS comuna,
    (result).lon AS longitud,
    (result).lat AS latitud
FROM
    geocode;

```

El resultado será el siguiente

direccion_norm	barrio	comuna	longitud	latitud
KR 20 13 A 45	LAS AMERICAS	COMUNA 1	-8602713.06642356	134491.159088312

Un ejemplo similar puede ejecutarse para la geocodificación de direcciones según la nomenclatura barrio manzana predio:

```

=# WITH geocode AS (
    SELECT geocode_address('Br La Paz', 'Manzana K Casa 9') AS result
)
SELECT
    (result).address AS direccion_normalizada,
    (result).neighborhood AS barrio,
    (result).commune AS comuna,
    (result).lon AS longitud,
    (result).lat AS latitud
FROM
    geocode;

```

El resultado será el siguiente

direccion_norm	barrio	comuna	longitud	latitud
MZ K CS 9	LA PAZ	COMUNA 4	-8600523.43231094	133201.206467672

## ANEXO C.

### Manual de usuario visor cartográfico PASTO VIEW

#### 1. Instalación

Como prerequisite de la instalación aplicativo web se debe tener instalado el geocodificador de direcciones urbanas, instalación descrita en el Anexo A.

Los pasos descritos a continuación fueron ejecutados sobre la plataforma de sistema operativo Linux/Ubuntu en su versión 14.04.

##### 1.1. Instalación Servidor de Aplicaciones Glassfish

Primero se descarga la versión 3.1.2 para GNU/linux desde el sitio web de Oracle y se ejecuta lo siguiente:

```
$ sh ogs-3.1.2.2-unix-ml.sh
```

Se debe descargar el driver JDBC de postgresql desde el sitio web de PostgreSQL y copiarlo en el directorio glassfish3/glassfish/domains/domain1/lib. Para iniciar el servidor se ejecuta lo siguiente:

```
$ ./glassfish3/glassfish/bin/startserv
```

Con esto en el navegador se ingresa con la dirección <http://localhost:4848>, y se ingresa a la consola de administración, se ingresa usuario y contraseña de haber escrito una en la instalación de glassfish.

Ir a “Resources/JDBC/Connection Pools” y crear una nueva conexión con los datos que muestra la Figura C1 y luego clic en siguiente.

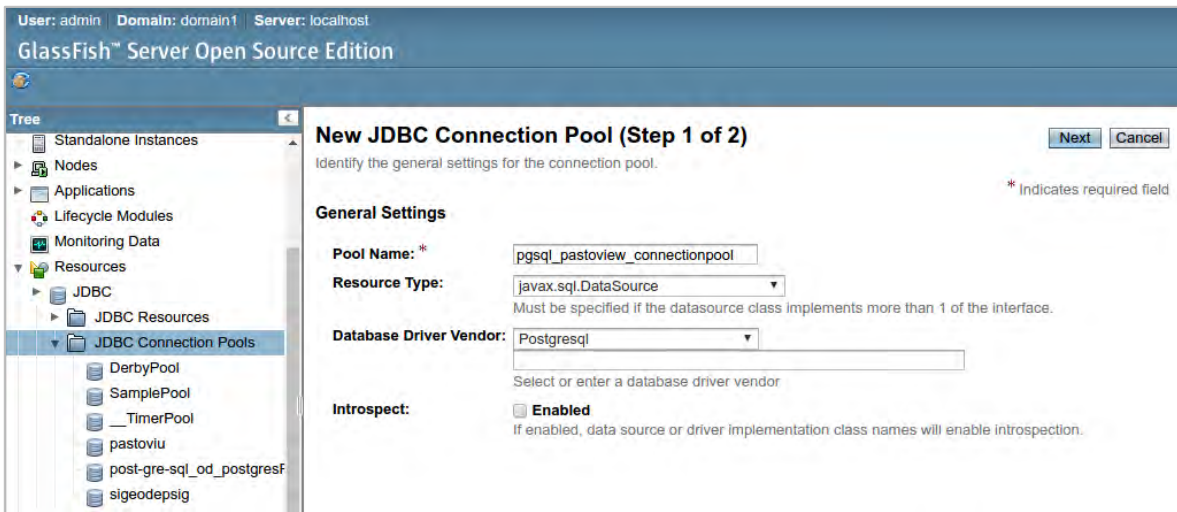


Figura C1. Creación nueva conexión a base de datos

Seleccione el origen de datos de nombre de clase org.postgresql.ds.PGSimpleDataSource y escribir las siguientes propiedades adicionales como muestra la Figura C2.

Additional Properties (5)			
Name	Value	Description:	
password	<contraseña_usuario_postgresql>		
user	<nombre_usuario_postgresql>		
port	5432		
host	localhost		
databaseName	geocoder		

Figura C2. Configuración propiedades de conexión.

Con esto se guarda las conexiones y se da clic en finalizar para guardar la conexión. Luego en “Resources/JDBC/JDBC Resources” se escribe en nombre JNDI y escoge el Pool Name creado anteriormente como lo muestra la Figura C3.

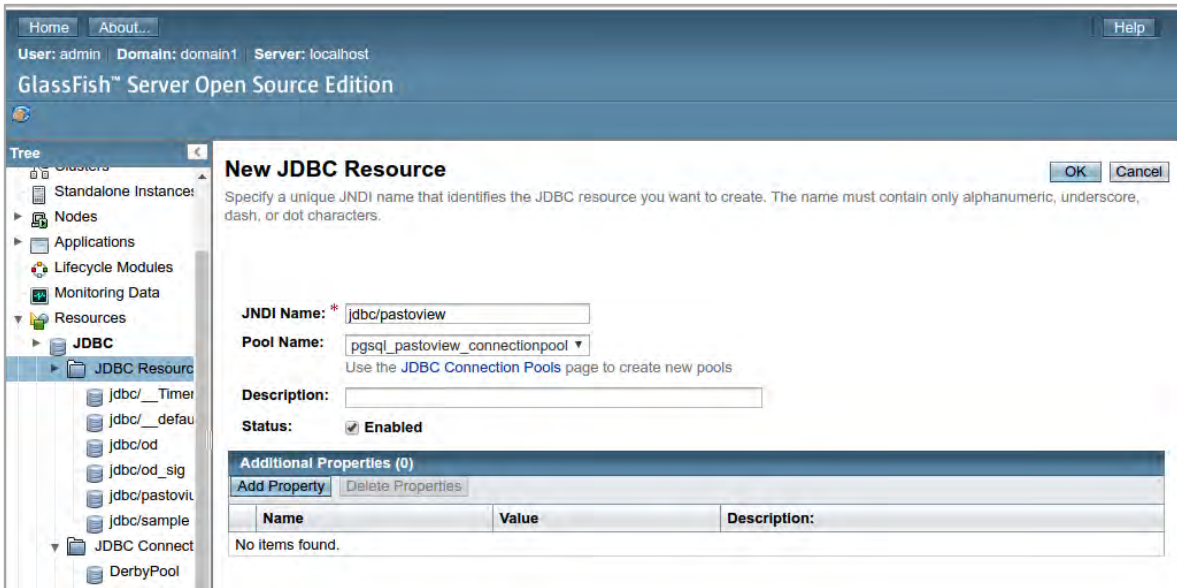


Figura C3. Creación recurso JDBC

Por último, en el menú Aplicaciones se selecciona la opción desplegar cargando el archivo “.war” donde se tiene almacenada la aplicación web como muestra la Figura C4. Si todo sale bien se direccionará a la página donde se muestra la url de la aplicación instalada. Al ingresar se desplegará la página principal de la aplicación como lo muestra la Figura C5.

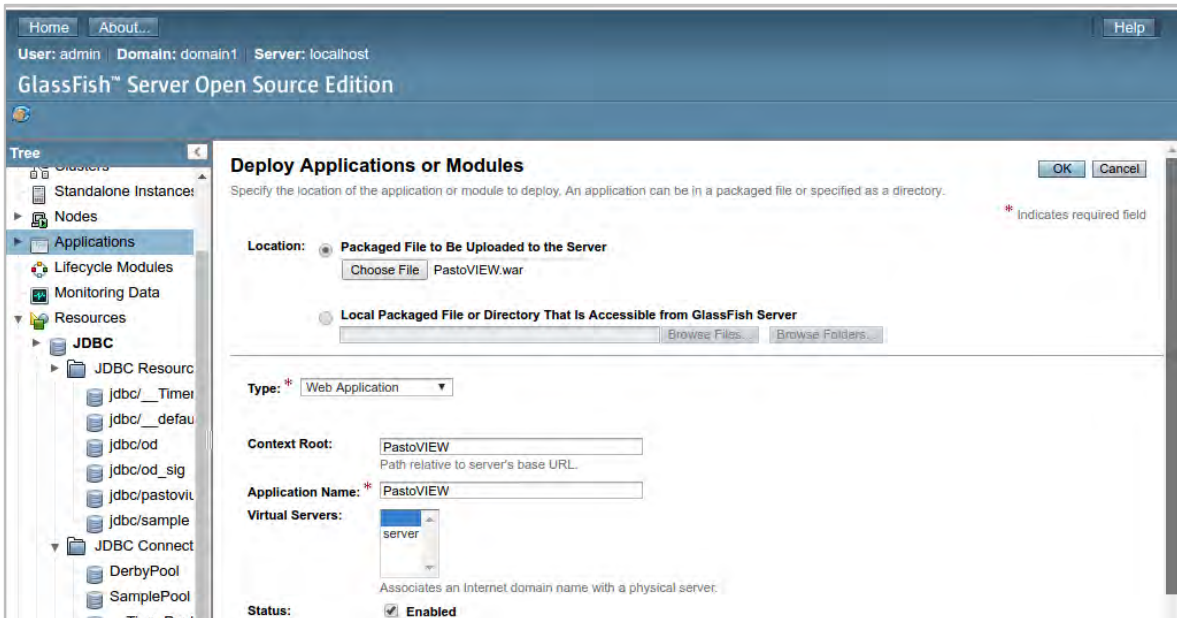


Figura C4. Carga del aplicativo web en el servidor glassfish

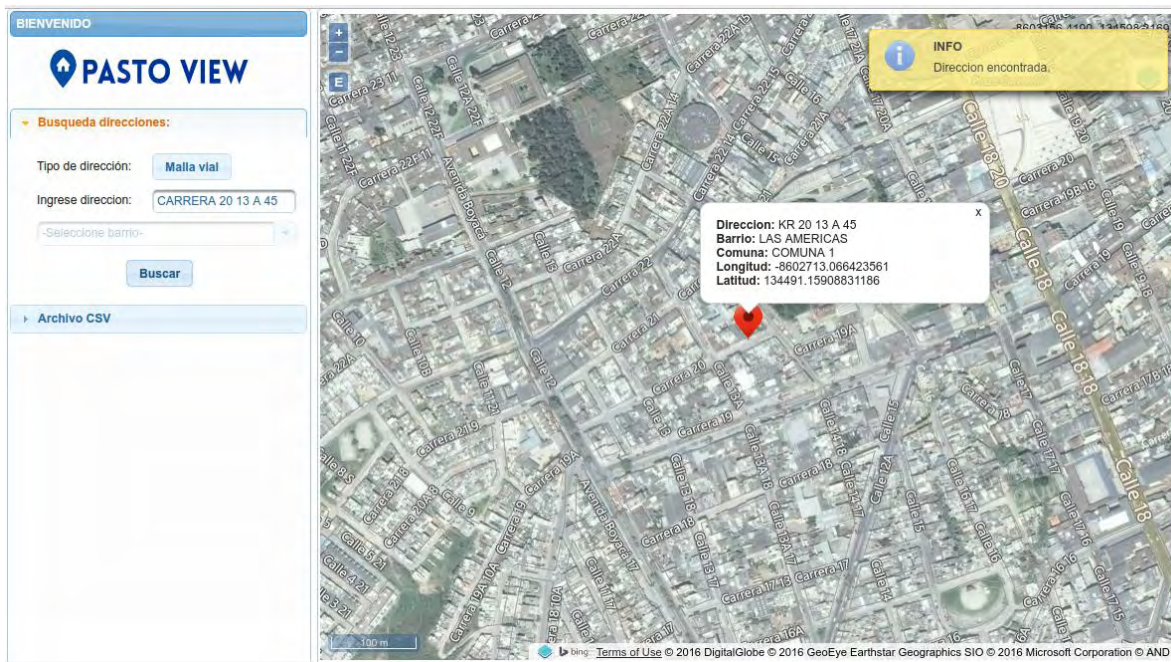


Figura C5. Visor PASTOVIEW en ejecución

## 2. Manual de usuario

El visor cartográfico PASTO VIEW es una extensión del geocodificador de direcciones urbanas para el municipio de Pasto. Este consiste en un aplicativo web permite al usuario acceder a la búsqueda de direcciones urbanas, procesamiento de archivos separados por comas CSV y la visualización de estos datos procesados, soportando la geocodificación de direcciones urbanas asignadas según la malla vial y según la nomenclatura barrio - manzana - predio.

### 2.1. Menú de PASTOVIEW

El acceso a las funcionalidades de búsqueda de direcciones urbanas y procesamiento de archivos CSV se encuentran en el menú principal del aplicativo web.

**2.1.1. Búsqueda direcciones:** A continuación se explicará la búsqueda de direcciones urbanas, opción que se muestra en la figura C6.

BIENVENIDO

**PASTO VIEW**

Busqueda direcciones

Tipo de dirección: **Malla vial**

Ingrese dirección:

-Seleccione barrio-

Buscar

Archivo CSV

Figura C6. Menú búsqueda de direcciones urbanas

En la primera opción denominada Tipo de dirección se selecciona el tipo de dirección que se desea consultar al geocodificador de direcciones. Al hacer clic sobre este botón, este intercambiará su valor según las dos opciones disponibles: Malla vial y Manzanas.

Si la opción Manzanas es seleccionada, se habilitará la opción Seleccionar barrio, campo requerido por el geocodificador para realizar la búsqueda de este tipo de direcciones urbanas. Esta configuración puede verse en la figura F7.

BIENVENIDO

**PASTO VIEW**

Busqueda direcciones

Tipo de dirección: **Manzanas**

Ingrese dirección:

LA PAZ

Buscar

Figura C7. Búsqueda de direcciones tipo Manzanas.

El campo Ingrese dirección se ingresa textualmente la dirección a buscar. Al ingresar la dirección que se desea consultar, se debe considerar los componentes mínimos de cada tipo de dirección.

Para el caso de malla vial debe ingresarse la vía principal, la vía generadora o secundaria y la distancia de la intersección de estas dos vías con la puerta de ingreso al predio buscado (ej. Carrera 20 13 A 45). Un ejemplo de esta búsqueda puede verse en la figura C8.

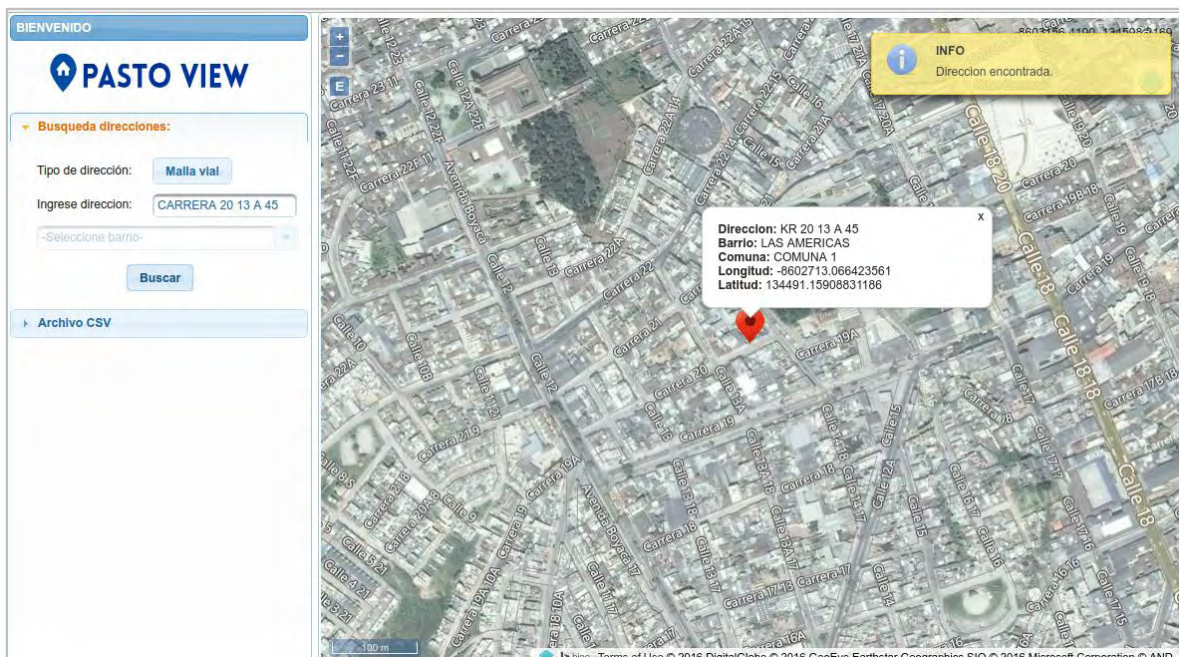


Figura C8. Búsqueda de dirección según la malla vial en PASTOVIEW

Si la opción seleccionada es manzanas, en el campo de texto solo debe ingresarse el identificador de manzana y el identificador de la casa o predio (ej. manzana k casa 9). Para este caso se habilita el campo Seleccione barrio, con el cual se especifica el barrio al que corresponde la nomenclatura de manzana ingresada.

Un ejemplo de esta búsqueda puede verse en la figura C9.

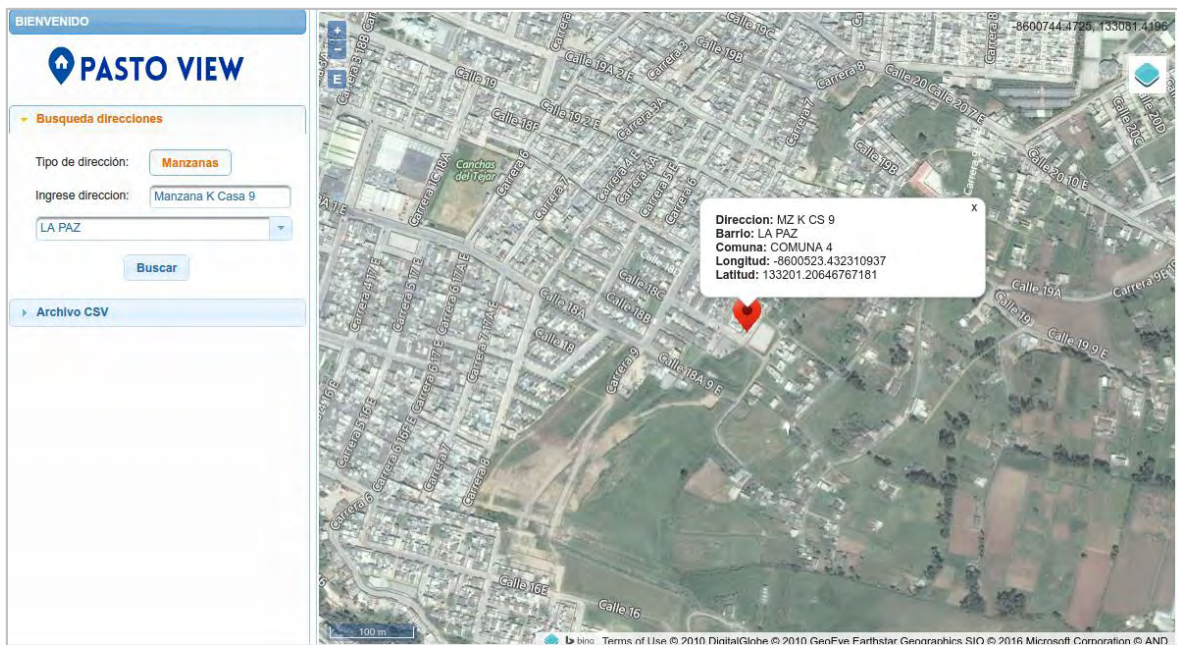


Figura C9. Búsqueda de dirección según la nomenclatura barrio - manzana predio en PASTOVIEW

Cualquiera sea el tipo de dirección seleccionado, si la dirección es encontrada en el repositorio de datos, se visualizará la ubicación urbana de la dirección mostrando información como la dirección estandarizada, el barrio y la comuna a la que pertenece y las coordenadas geográficas latitud/longitud encontradas.

**2.1.2. Archivos CSV:** A continuación se explica la funcionalidad subida y procesamiento de archivos CSV. Su implementación puede observarse en la figura C10.

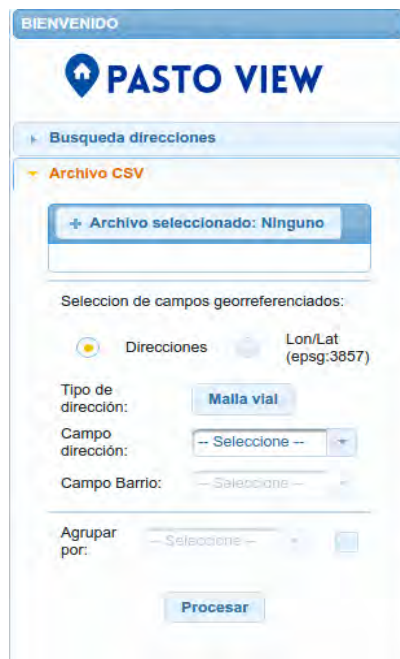


Figura C10. Funcionalidad de subida y procesamiento de datos CSV implementada en PASTO VIEW

El panel Archivos CSV se compone de 3 opciones principales:

- **Subir archivo CSV:** En esta opción se selecciona el archivo que será procesado por el visor cartográfico. Las condiciones que debe cumplir este archivo son:
  - ✓ Los datos almacenados en este archivo deben utilizar el carácter coma (,) como delimitador.
  - ✓ Debe tener un registro cabecera (header) que identifique los nombres de cada campo.
  - ✓ Este debe tener la extensión .CSV, de lo contrario se mostrara un mensaje de archivo inválido.

El archivo será subido al servidor de la aplicación correctamente si en la etiqueta Archivo subido se muestra el nombre del archivo seleccionado.. La figura C11 muestra un ejemplo de esta subida de archivos. De igual forma se cargará las listas de selección de los campos restantes.



Figura C11. Subida correcta e incorrecta de archivos CSV.

- **Selección de campos georreferenciados:** Esta opción permite seleccionar los campos del archivo CSV donde se encuentran los datos georreferenciados.

Un primer tipo de campo es Direcciones, aquí se especifica los campos tipo texto donde estén almacenados direcciones urbanas, según sea el caso, Malla vial y nomenclatura barrio - manzana – predio (Manzanas), que estén configuradas dentro de los datos CSV. Un ejemplo de esta configuración puede observarse en la figura C12.

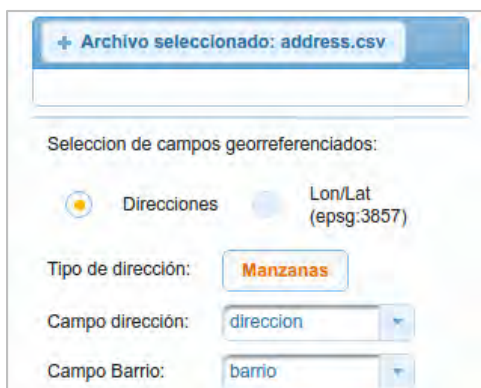


Figura C12. Configuración del archivo ‘address.csv’, el cual contiene direcciones asociadas al caso Manzanas.

Un segundo tipo de campo es Longitud/Latitud, aquí se especifica dos campos numéricos con decimales que corresponden a las coordenadas geográficas longitud/latitud asociada a ese registro. Estas coordenadas deben estar calculadas según el sistema de referencia espacial EPSG:3857. Un ejemplo de esta configuración puede observarse en la figura C13.



Figura C12. Configuración del archivo ‘cluster-fatales.csv’, seleccionando los campos ‘longitud’ y ‘latitud’ presentes en este archivo.

- **Agrupar por:** Esta opción por defecto aparecerá desactivada por lo cual los datos procesados se visualizarán georreferenciados a nivel de direcciones urbanas. La figura C13 muestra el procesamiento del archivo ‘clusters-fatales.csv’, ubicando cada registro procesado en su correspondiente ubicación.

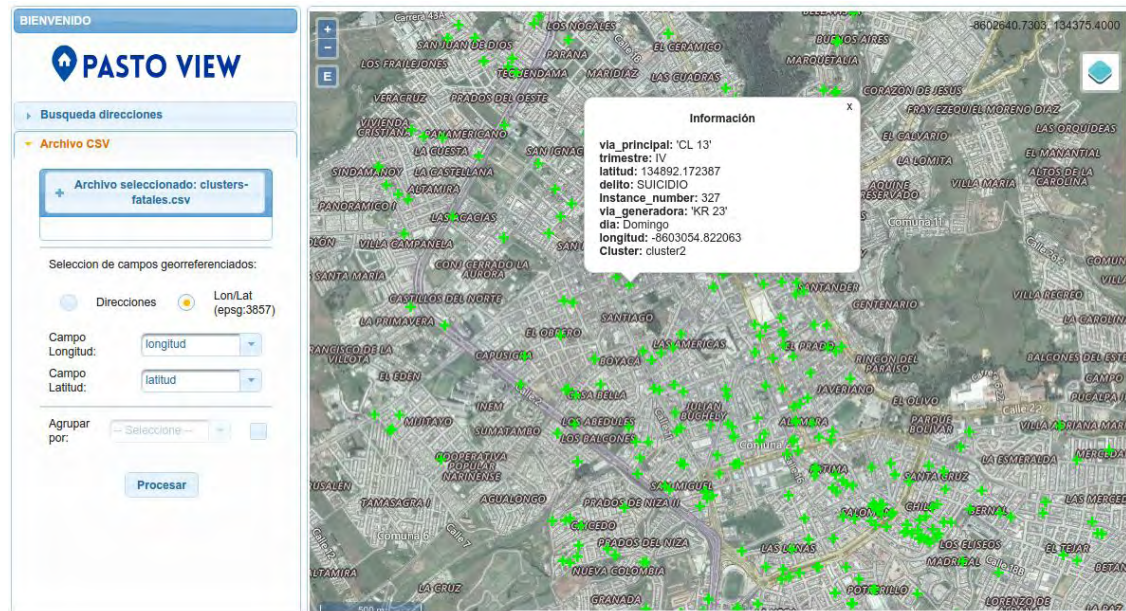


Figura C13. Procesamiento del archivo ‘clusters-fatales.csv’ sin agrupar.

Si esta opción se habilita, el visor cartográfico creará capas o mapas de visualización de acuerdo al campo seleccionado. Estas capas agruparán los datos según la característica o campo seleccionado, permitiendo deshabilitar una o varias de estas capas. La figura C14 muestra un ejemplo de esta configuración, en donde se agrupan los datos por el campo ‘Cluster’ deshabilitando las capas que corresponden a los valores ‘cluster0’ y ‘cluster1’.

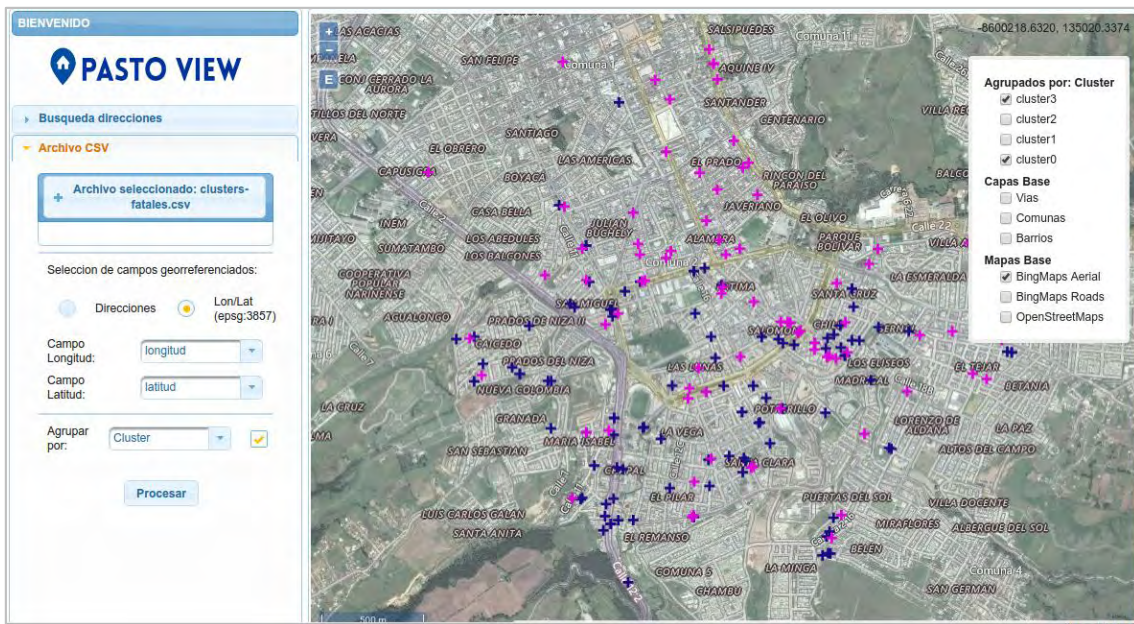


Figura C14. Procesamiento del archivo ‘clusters-fatales.csv’. Los datos se agrupan según los valores del campo ‘Clusters’.

El usuario podrá visualizar a los datos asociados a cada evento o registro procesado haciendo clic en cada punto. Al hacerlo se desplegará un dialogo con cada característica asociada a dicho punto o registro.

## ANEXO D.

### Análisis UML visor cartográfico PASTO VIEW

A continuación se describen aspectos del análisis realizado para definir las funcionalidades del visor cartográfico Pasto View.

- 1. Actores del Sistema:** El visor Pasto View tendrá acceso sin restricciones a la consulta de direcciones urbanas a cualquier usuario que ingrese al sistema. Los usuarios podrán buscar direcciones urbanas, visualizando la ubicación espacial de las mismas.
- 2. Funciones del sistema:** La tabla C1 muestra los requerimientos definidos para el visor cartográfico SIGEODEP SIG.

<b>Búsqueda de direcciones urbanas</b>		
<b>Ref.</b>	<b>Función</b>	<b>Descripción</b>
R1	Conexión con el geocodificador de direcciones urbanas.	El sistema debe conectarse al geocodificador de direcciones urbanas instalado en el gestor de base de datos PostgreSQL para realizar consultas.
R1.1	Búsqueda de direcciones urbanas según malla vial.	El sistema debe permitir al usuario consultar direcciones urbanas asignadas según la malla vial.
R1.2	Búsqueda de direcciones urbanas según nomenclatura barrio manzana predio.	El sistema debe permitir al usuario consultar direcciones urbanas asignadas según la nomenclatura barrio manzana predio.
R2	Visualización de direcciones encontradas.	El sistema debe permitir al usuario visualizar la ubicación de la dirección urbana sobre el mapa del municipio de Pasto sobre mapas temáticos.
R2.1	Visualización de información relacionada con la dirección encontrada	El sistema debe permitir al usuario visualizar la información urbana relacionada con la dirección urbana encontrada.

R3	Carga de cartográfica correspondiente al municipio de Pasto.	El sistema debe cargar el mapa satelital del municipio de Pasto que permita visualizar las calles y vías con aproximación a la realidad.
R4	Navegación sobre los mapas cargados.	El sistema debe permitir al usuario navegar sobre los mapas temáticos cargados (ampliar, reducir la visibilidad del mapa, activación/desactivación de capas),
<b>Archivos CSV</b>		
R5	Procesar archivos CSV	El sistema debe permitir procesar archivos separados por comas CSV con el fin de georreferenciar los registros que se encuentren almacenados.
R6	Seleccionar campos georreferenciados	El sistema debe permitir seleccionar el tipo de campos georreferenciados, los cuales pueden ser direcciones urbanas y/o coordenadas geográficas.
R6.1	Procesamiento de campo dirección urbana y barrio	El sistema debe permitir identificar el tipo de dirección urbana (Malla vial o Manzanas) presente entre los datos CSV y seleccionar los campos relacionados con direcciones urbanas, estos son dirección urbana y barrio.
R6.2	Procesamiento de coordenadas geográficas Longitud/Latitud	El sistema debe permitir seleccionar los campos donde se encuentren las coordenadas geográficas Longitud y Latitud. Estas coordenadas deben estar configuradas según el sistema de referencia espacial 3857.
R7	Agrupar datos	El sistema debe permitir agrupar los datos procesados de acuerdo a la selección de un determinado campo de datos.
R7.1	Visualizar capas de datos agrupados	El sistema debe permitir crear capas o mapas de visualización de datos geocodificados, de acuerdo a la configuración de agrupación de datos que se haya seleccionado.

**3. Casos de uso:** Un caso de uso es una secuencia de interacciones que se desarrollarán entre un sistema y sus actores en respuesta a un evento que inicia un actor principal sobre el propio sistema.

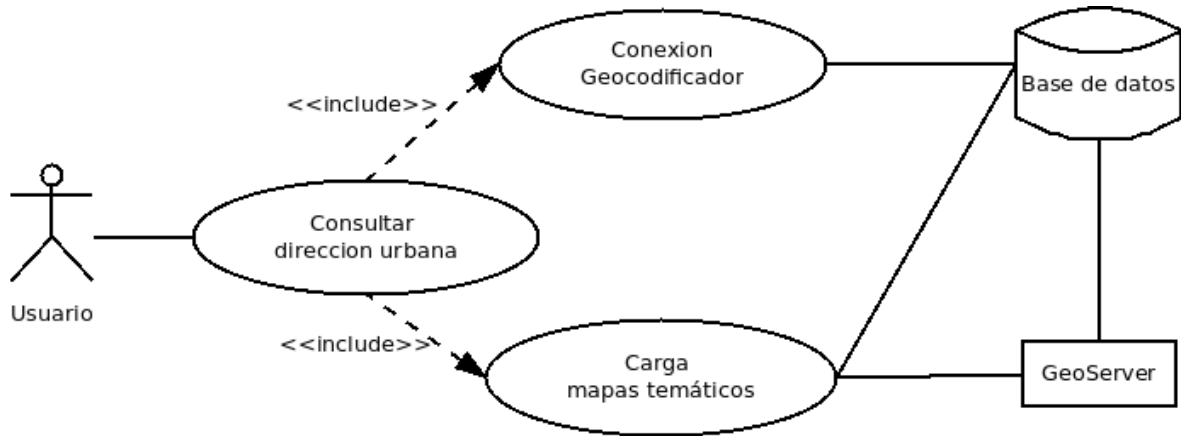
La funcionalidad del visor Pasto View se describe en dos casos de uso

- Consulta de direcciones urbanas.
  - Procesamiento de archivos separados por comas CSV.
- Caso de uso 1. Consulta de direcciones urbanas.

Caso de uso:	Consulta de direcciones urbanas
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Geocodificador de direcciones urbanas instalado.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> </ul>
Post-condiciones:	Dirección urbana ubicada en el área urbana del municipio de Pasto de acuerdo a las coordenadas geográficas latitud/longitud retornadas por el geocodificador.
Flujo Básico	
<ol style="list-style-type: none"> <li>1. El sistema despliega los formularios de ingreso de direcciones urbanas según la malla vial y según la nomenclatura barrio – manzana – predio.</li> <li>2. El usuario ingresa una dirección urbana en el formulario correspondiente y selecciona la opción Buscar.</li> <li>3. El sistema captura la dirección urbana ingresada generando y ejecutando la consulta SQL que llama al geocodificador de direcciones urbanas.</li> <li>4. El sistema recupera los datos arrojados por el geocodificador de direcciones.</li> <li>5. El sistema crea un punto en el mapa temático de acuerdo a las coordenadas latitud longitud recibida.</li> <li>6. El sistema muestra la información urbana de la dirección consultada sobre el punto ubicado.</li> <li>7. El sistema muestra un mensaje informando al usuario que la dirección ingresada fue encontrada.</li> </ol>	
Flujo alternativo	
<ol style="list-style-type: none"> <li>4a. El sistema verifica que el geocodificador no obtuvo resultados asociados a la dirección consultada.</li> <li>4b. El sistema muestra un mensaje de error informando que la dirección ingresada no fue encontrada.</li> </ol>	

La figura D1 muestra el diagrama de caso de uso.

Figura D1. Diagrama de caso de uso Consulta de direcciones urbanas



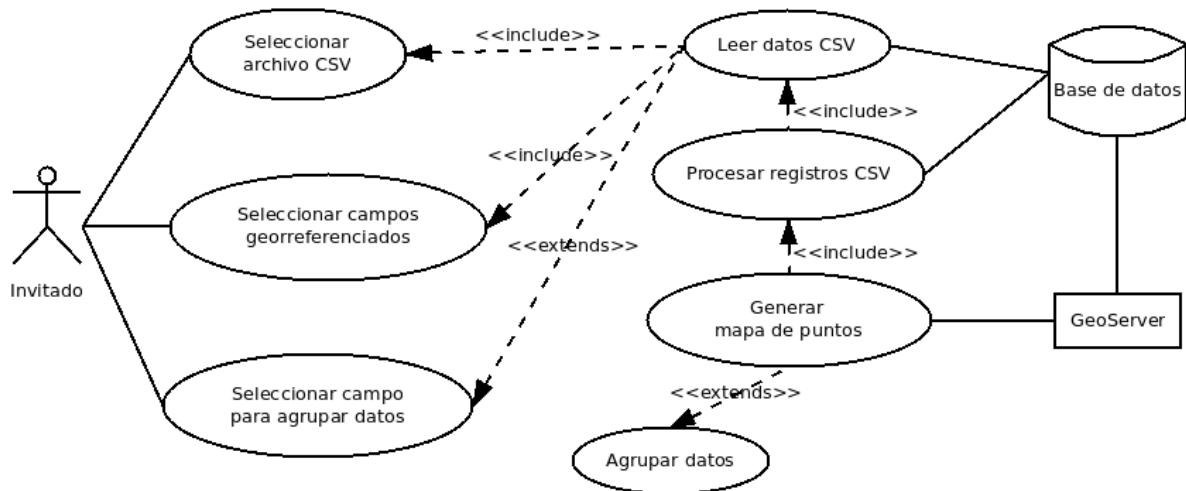
- Caso de uso Procesamiento de archivos CSV.

Caso de uso:	Procesamiento de archivos CSV
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Geocodificador de direcciones urbanas instalado.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> </ul>
Post-condiciones:	Mapa(s) de puntos asociado(s) al procesamiento de los registros encontrados en el archivo CSV.
Flujo Básico	
<ol style="list-style-type: none"> <li>1. El sistema despliega el formulario de selección y procesamiento de archivos CSV.</li> <li>2. El usuario selecciona la opción 'Seleccionar archivo' y buscar el archivo que desea procesar.</li> <li>3. El sistema procesa el archivo seleccionado cargando los nombres de cada campo encontrado en las listas de selección de campos.</li> <li>4. El usuario configura las listas de campos de acuerdo a la estructura del archivo CSV y presiona el botón 'Procesar'.</li> <li>5. El sistema procesa cada registro de acuerdo a los campos georreferenciados seleccionados.</li> <li>6. El sistema despliega la capa de puntos con cada evento o registro procesado informando a usuario la cantidad de registros procesados exitosamente.</li> </ol>	
Flujo alternativo	
<ol style="list-style-type: none"> <li>2a. El sistema verifica que el archivo no tiene la extensión .csv.</li> <li>2b. El sistema informa al usuario que el archivo no es válido.</li> <li>4a. Si la opción Agrupar datos es habilitada, el usuario debe seleccionar el campo por el cual se agruparan los datos.</li> </ol>	

- 4b. El sistema procesa cada registro de acuerdo a los campos georreferenciados seleccionados.
- 4c. El sistema despliega la capa de puntos con cada evento o registro procesado informando a usuario la cantidad de registros procesados exitosamente.

La figura D2 muestra el diagrama de caso de uso.

Figura D2. Diagrama de caso de uso Procesamiento de archivos CSV



## ANEXO E.

### Procedimientos SQL de geocodificación de eventos delictivos

#### 1. Geocodificación de eventos delictivos fatales.

```
--creación de la estructura de datos
CREATE TABLE
    geocoded_fatal_injuries AS(
        SELECT
            fatal_injury_id,
            injury_address,
            NULL::varchar AS address,
            NULL::varchar AS neighborhood,
            NULL::varchar AS commune,
            NULL::double precision AS lon,
            NULL::double precision AS lat
        FROM
            fatal_injuries
        WHERE
            injury_address IS NOT NULL AND
            injury_address NOT LIKE 'SIN DATO'
    );
--Insercion (update) de campos vacios faltantes utilizando el
--geocoder de direcciones urbanas según malla vial
WITH geocode_injuries AS(
    SELECT
        fatal_injury_id AS id,
        injury_address,
        --Aqui se llama al geocoder
        geocode_address(injury_address) AS geocode
    FROM
        fatal_injuries
) UPDATE
    geocoded_fatal_injuries
SET
    address = (geocode).address,
    neighborhood = (geocode).neighborhood,
    commune = (geocode).commune,
    lon = (geocode).lon,
    lat = (geocode).lat
FROM
    geocode_injuries
```

```

WHERE
    (geocode).address IS NOT NULL AND
    fatal_injury_id = id;

--inserción (update) de campos vacíos faltantes utilizando el geocoder
--según nomenclatura barrio-manzana-predio
WITH geocode_injuries AS (
    SELECT
        fatal_injury_id AS id,
        injury_address,
        neighborhood_name,
        --Aqui se llama al geocoder
        geocode_address(neighborhood_name, injury_address) AS
geocode
    FROM
        fatal_injuries
    JOIN neighborhoods ON injury_neighborhood_id=neighborhood_id
    WHERE
        injury_address NOT LIKE 'SIN DATO' AND
        injury_address IS NOT NULL
) UPDATE
    geocoded_fatal_injuries
SET
    address = (geocode).address,
    neighborhood = (geocode).neighborhood,
    commune = (geocode).commune,
    lon = (geocode).lon,
    lat = (geocode).lat
FROM
    geocode_injuries
WHERE
    (geocode).address IS NOT NULL AND
    fatal_injury_id = id;

```

## 2. Geocodificación de eventos delictivos no fatales.

```

--creación de la estructura de datos
CREATE TABLE
    geocoded_non_fatal_injuries AS (
        SELECT
            non_fatal_injury_id,
            injury_address,
            NULL::varchar AS address,
            NULL::varchar AS neighborhood,
            NULL::varchar AS commune,
            NULL::double precision AS lon,
            NULL::double precision AS lat

```

```

        FROM
            fatal_injuries
        WHERE
            injury_address IS NOT NULL AND
            injury_address NOT LIKE 'SIN DATO'
    );
--Insercion (update) de campos vacios faltantes utilizando el
--geocoder de direcciones urbanas según malla vial
WITH geocode_injuries AS(
    SELECT
        non_fatal_injury_id AS id,
        injury_address,
        --Aqui se llama al geocoder
        geocode_address(injury_address) AS geocode
    FROM
        non_fatal_injuries
)UPDATE
    geocoded_non_fatal_injuries
SET
    address = (geocode).address,
    neighborhood = (geocode).neighborhood,
    commune = (geocode).commune,
    lon = (geocode).lon,
    lat = (geocode).lat
FROM
    geocode_injuries
WHERE
    (geocode).address IS NOT NULL AND
    non_fatal_injury_id = id;

--inserción (update) de campos vacíos faltantes utilizando el geocoder
--según nomenclatura barrio-manzana-predio
WITH geocode_injuries AS (
    SELECT
        non_fatal_injury_id AS id,
        injury_address,
        neighborhood_name,
        --Aqui se llama al geocoder
        geocode_address(neighborhood_name, injury_address) AS
geocode
    FROM
        non_fatal_injuries
    JOIN neighborhoods ON injury_neighborhood_id=neighborhood_id
    WHERE
        injury_address NOT LIKE 'SIN DATO' AND
        injury_address IS NOT NULL
)UPDATE
    geocoded_fatal_injuries
SET
    address = (geocode).address,

```

```
neighborhood = (geocode).neighborhood,  
commune = (geocode).commune,  
lon = (geocode).lon,  
lat = (geocode).lat  
FROM  
geocode_injuries  
WHERE  
(geocode).address IS NOT NULL AND fatal_injury_id = id;
```

## ANEXO F.

### Manual de usuario visor cartográfico SIGEODEP SIG

#### 1. Instalación

A continuación se describen el procedimiento para instalar desde cero el visor cartográfico SIGEODEP SIG, que a diferencia de PASTO VIEW, este aplicativo web debe conectarse directamente con la base de datos con eventos delictivos del observatorio del delito del municipio de Pasto. Estos pasos fueron ejecutados sobre la plataforma de sistema operativo Linux/Ubuntu en su versión 14.04.

##### 1.1. Gestor de bases de datos PostgreSQL con extensión PostGIS

Instalación librerías requeridas para el funcionamiento de PostgreSQL:

```
$ sudo apt-get install libxml2 libxml2-dev libproj-dev libgeos-3.4.2  
libgdal-dev libjson0 libjson0-dev
```

Instalación del gestor PostgreSQL 9.4 - PostGIS 2.1 - Pgadmin3:

```
$ sudo apt-get install postgresql-9.4 postgresql-9.4-postgis-2.1  
postgresql-client-9.4 postgresql-server-dev-9.4 postgresql-contrib-9.4  
pgadmin3
```

Configuración de la contraseña del usuario postgres:

```
$ su postgres  
$ psql postgres
```

```
=# ALTER USER 'postgres' WITH PASSWORD 'postgres1'  
=# \q
```

## 1.2. Extensión para PostgreSQL pg\_similarity

Fue utilizado el script jarowinkler de la extensión pg\_similarity, permitiendo calcular la similitud entre dos cadenas de texto. Esta funcionalidad es utilizada en la estandarización del barrio ingresado junto con la dirección asignada según la nomenclatura barrio – manzana - predio.

Instalación:

```
$ cd pg_similarity
$ $EDITOR Makefile
$ USE_PGXS=1 make
$ USE_PGXS=1 make install
```

## 1.3. Restauración base de datos

Con las siguientes instrucciones se procede a crear y restaurar la base de datos de delitos.

```
$ su postgres
$ psql postgres
postgres=# CREATE DATABASE od;
postgres=# \q
$ psql od
od=# \i /ruta/archive/sql/od.sql
```

## 1.4. Instalacion servidor de aplicaciones Glassfish

Se descarga la versión 3.1.2 para GNU/linux desde el sitio web de Oracle y se ejecuta lo siguiente

```
$ sh ogs-3.1.2.2-unix-m1.sh
```

Se debe descargar el driver JDBC de postgresql desde el sitio web de PostgreSQL y copiarlo en el directorio glassfish3/glassfish/domains/domain1/lib. Para iniciar el servidor se ejecuta lo siguiente:

```
$ ./glassfish3/glassfish/bin/startserv
```

Con esto en el navegador se ingresa con la dirección <http://localhost:8080>, y se ingresa a la consola de administración, se ingresa usuario y contraseña de haber escrito una en la instalación de glassfish.

Ir a “Resources/JDBC/Connection Pools” y crear una nueva conexión con los datos que muestra la Figura F1 y luego clic en siguiente.

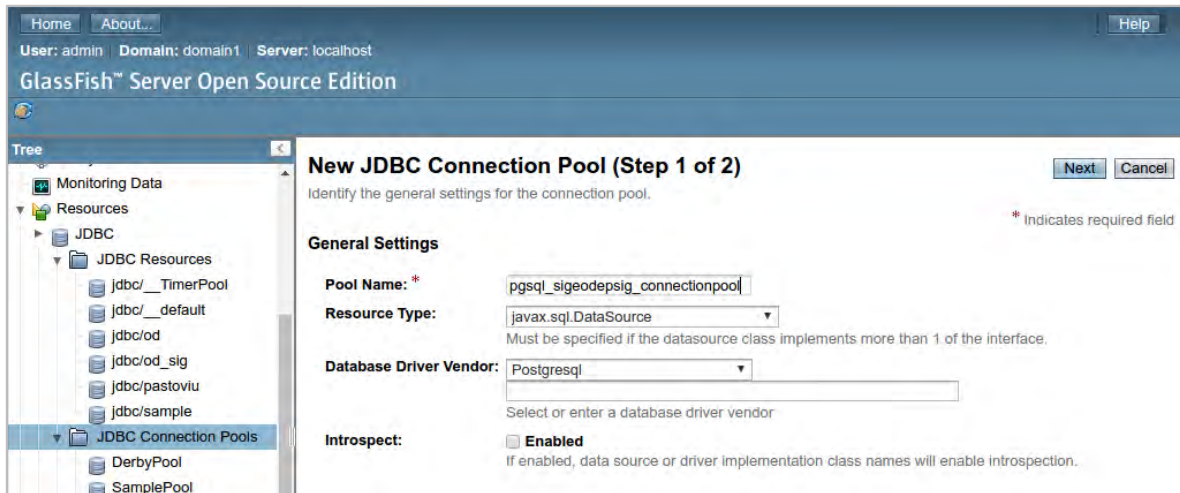


Figura F1. Creación nueva conexión a base de datos

Seleccione el origen de datos de nombre de clase org.postgresql.ds.PGSimpleDataSource y escribir las siguientes propiedades adicionales como muestra la Figura F2.

Additional Properties (7)			
Name	Value	Description:	
serverName	localhost		
User	<nombre_usuario_postgresql>		
databaseName	od		
portNumber	5432		
Password	<contraseña_usuario_postgresql>		
driverClass	org.postgresql.Driver		
URL	jdbc:postgresql://localhost:5432/od		

Figura F2. Configuración propiedades de conexión.

Con esto se guarda las conexiones y se da clic en finalizar para guardar la conexión. Luego en “Resources/JDBC/JDBC Resources” se escribe en nombre JNDI y escoge el Pool Name creado anteriormente como lo muestra la Figura F3.

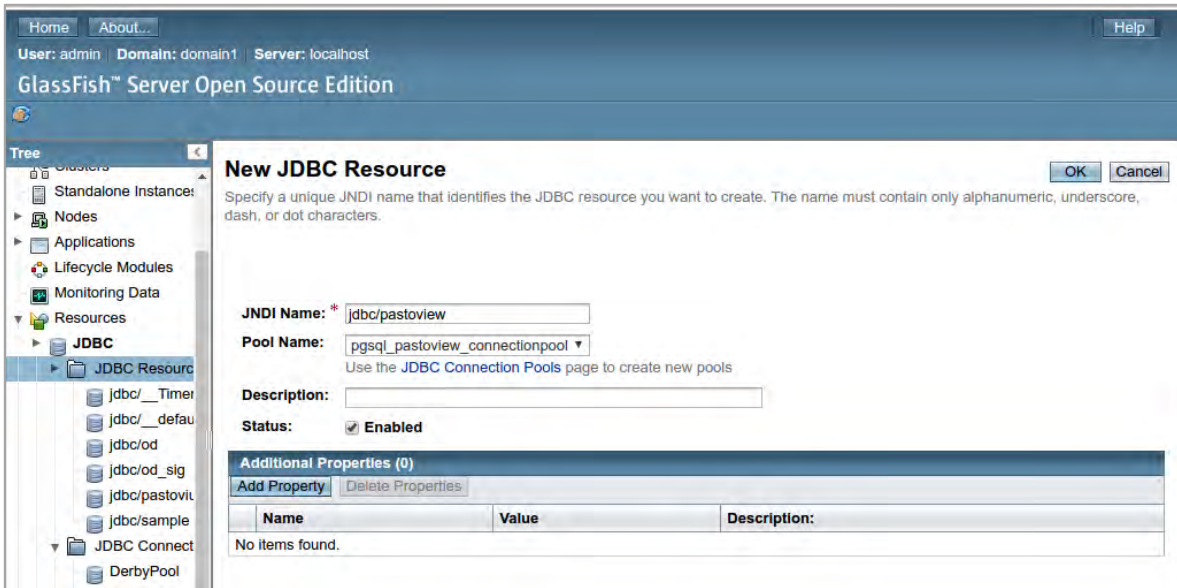


Figura F3. Creación recurso JDBC

Por último, en el menú Aplicaciones se selecciona la opción desplegar cargando el archivo “.war” donde se tiene almacenada la aplicación web como muestra la Figura D4. Si todo sale bien se redireccionará a una página donde se muestra la url de la aplicación instalada. Al ingresar se desplegará la página principal de la aplicación como lo muestra la Figura F5.

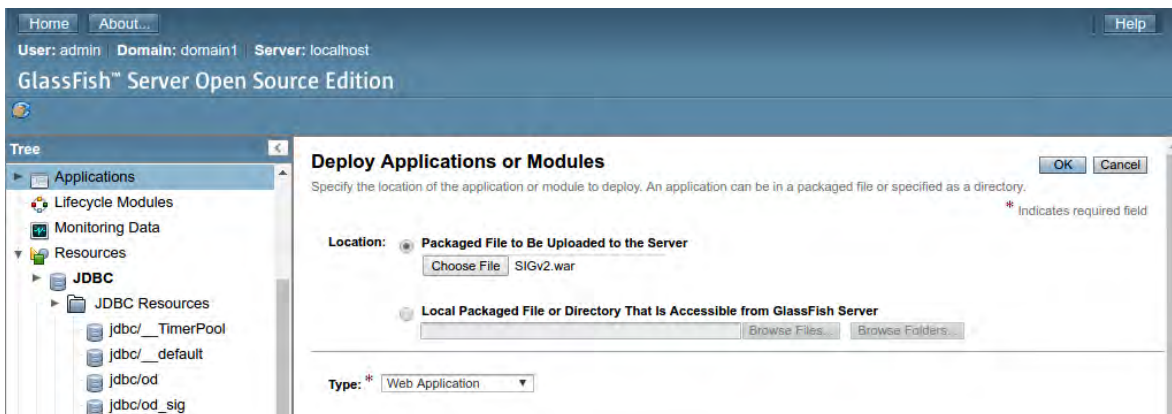


Figura F4. Carga del aplicativo web al servidor glassfish



Figura F5. Visor SIGEODEP SIG en ejecución

## 2. Manual de usuario

El visor cartográfico SIGEODEP SIG tiene como objetivo geocodificar eventos delictivos fatales y no fatales registrados por el observatorio del delito del municipio de Pasto a nivel de direcciones urbanas para identificar patrones visuales delictivos utilizando mapas de calor.

### 2.1. Ingreso a la herramienta

Al dar clic en el botón Ingresar se ingresa a la página principal del aplicativo web, permitiendo al usuario acceder a las funcionalidades implementadas. La figura F6 muestra esta página principal.

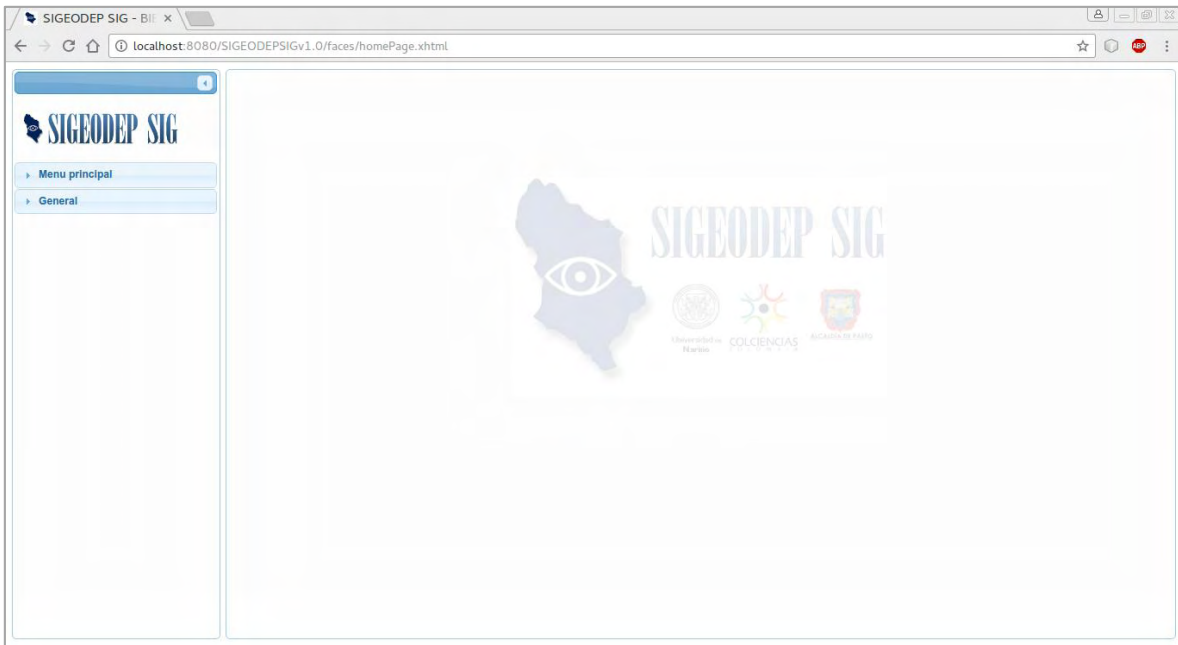


Figura F6. Página principal SIGEODEP SIG

La página principal se compone de un menú principal donde se encuentran los enlaces a las páginas con las funcionalidades implementadas y un menú general con información de la sesión iniciada y la opción de cerrar. La figura F7 muestra en detalle estas opciones.



Figura F7. Menú principal y opciones generales disponibles en la página principal de SIGEODEP SIG

## 2.2. Lesiones de Causa Externa

En esta página el usuario tendrá acceso a la funcionalidad de creación de mapas temáticos a partir del cruce y configuración de variables y el conteo de eventos delictivos en zonas específicas o de interés en el área urbana del municipio de Pasto. La figura F8 muestra la implementación de esta página.

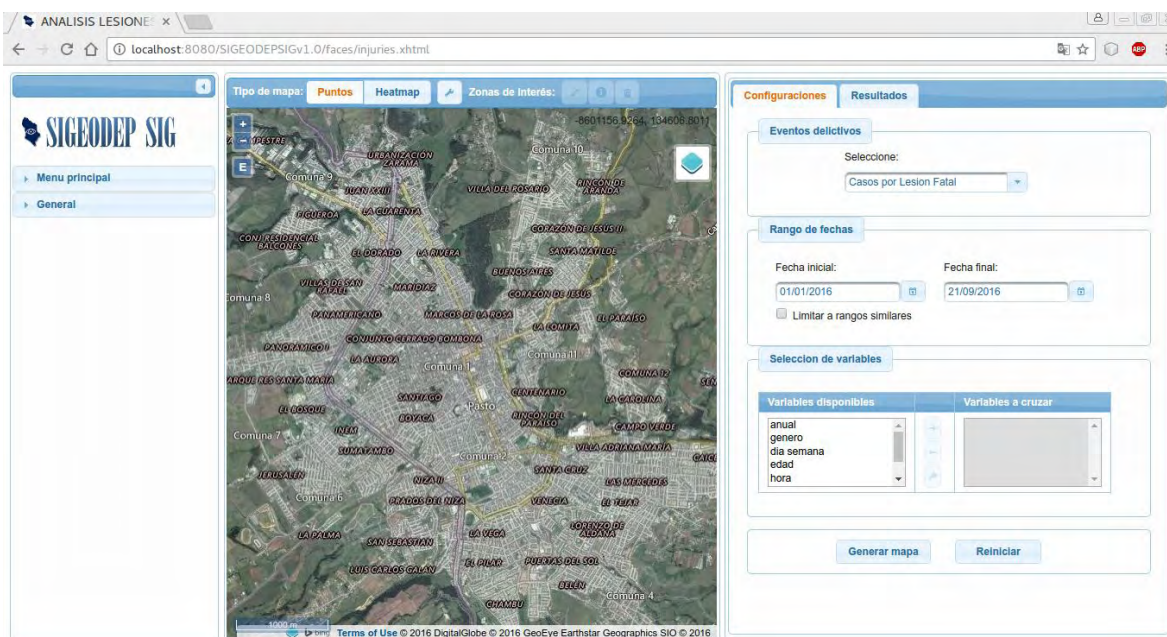


Figura F8. Lesiones de Causa Externa, pagina web principal.

### 2.2.1. Panel de Configuraciones/Resultados

Este panel ubicado en el lado derecho del aplicativo (véase figura F9). Un panel principal denominado ‘Configuraciones’ permite acceder a la configuración de atributos sobre los cuales se generaran los mapas temáticos con los eventos delictivos georreferenciados a nivel de direcciones urbanas. El panel auxiliar denominado ‘Resultados’ permite visualizar los resultados arrojados por la consulta de zonas de interés a partir de los mapas de calor generados por la aplicación. El uso de este panel se explicará más adelante.

En el panel de configuraciones se identifican tres opciones principales:

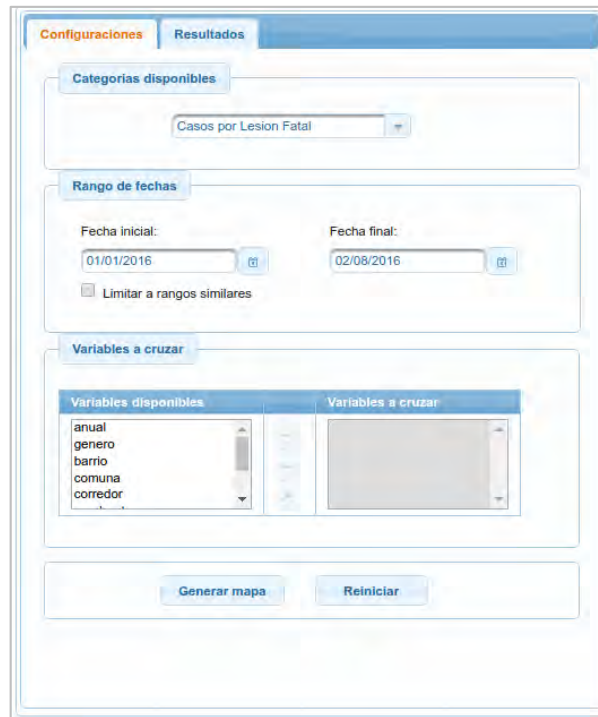


Figura F9. Panel de opciones de configuración de mapas temáticos

**Categorías disponibles:** Aquí se enlistan los eventos delictivos fatales y no fatales disponibles para ser georreferenciados a nivel de direcciones urbanas (Véase figura F10).

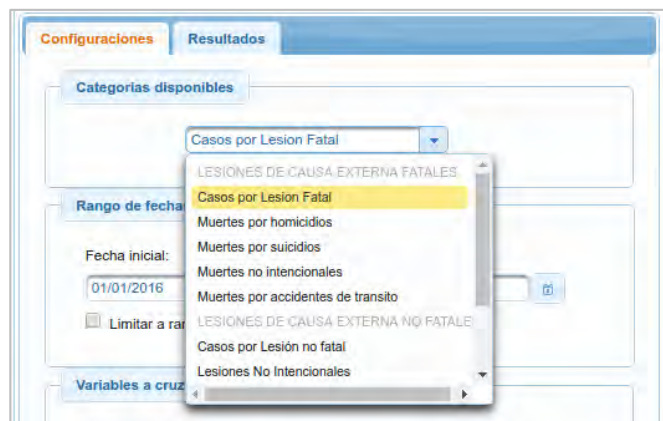


Figura F10. Categorías o tipos de eventos delictivos disponibles

**Rango de fechas:** Aquí se establece el rango o intervalo de fechas entre los cuales los eventos delictivos deben estar ubicados de acuerdo a su fecha de ocurrencia (véase figura F11).

Figura F11. Configuración de rango de fechas inicial y final

- **Limitar a rangos similares:** Esta opción permite que los datos que se tomen de cada año tengan un mismo rango de fechas existente, para explicarlo se realizara un ejemplo:

Cuando deseamos sacar el número de homicidios por año desde el 01/01/2012 hasta el 01/03/2014. Si no habilitamos la opción limitar a rangos similares obtenemos:

Casos de homicidios por año de 2011 a 2014 sin limitar a rangos similares				
Año	2011	2012	2013	2014
Número de Casos	100	110	120	20
Datos se obtuvieron de	Ene 2011 A Dic 2011	Ene 2012 A Dic 2012	Ene 2013 A Dic 2013	Ene 2014 A Marzo 2011

Lo importante a considerar es que del 2011 al 2013 se tomaron los datos de enero a diciembre, pero en el caso del año 2014 se tomaron los datos de enero a Marzo (Tres meses). Ahora bien la utilidad de limitar a rangos similares radica en que los datos del 2014 solo los tenemos los homicidios ocurridos hasta marzo, entonces los demás años no deberían tomar el total del año completo sino al igual que 2014 tomar solo los datos que van de enero a marzo de cada año, por tanto cuando habilitamos la opción limitar a rangos similares obtenemos:

Casos de homicidios por año de 2011 a 2014 limitando a rangos similares				
Año	2011	2012	2013	2014
Número de Casos	24	15	12	20
Datos se obtuvieron de	Ene 2011 A Marzo 2011	Ene 2012 A Marzo 2012	Ene 2013 A Marzo 2013	Ene 2014 A Marzo 2011

Se observa que los datos disminuyen para los tres primeros años por que ahora el rango de cada año va desde enero a marzo en cada uno de los años.

La utilidad de esta opción es que permite realizar un comparativo según los datos que el observatorio tenga hasta un determinado momento.

**Selección de variables:** Aquí se establece las variables que serán cruzadas para generar los mapas temáticos (véase figura F12).



Figura F12. Selección y configuración de variables.

**Variables disponibles:** El listado de variables disponibles nos muestra cuales son las variables que el sistema permite adicionar al proceso para la creación de los mapas temáticos. Por cada tipo de evento delictivo dispuesto en la sección Categorías disponibles tiene su propio listado de 'variables disponibles' (véase figura F13).

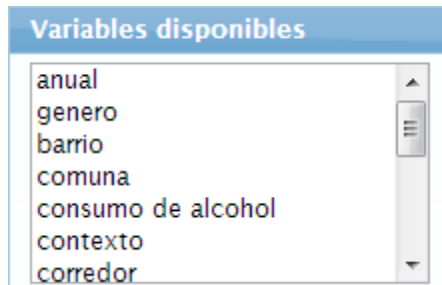


Figura F13. Variables asociadas a una categoría de evento delictivo.

**Variables a cruzar:** El listado de variables a cruzar son las variables que el usuario desea procesar según la categoría de evento delictivo seleccionada (véase figura F14). Los diferentes cruces de variables soportan un máximo de tres variables a cruzar, (se explica más detalladamente en la sección: 'Mínimo y máximo número de variables del cruce' que se explica más adelante).

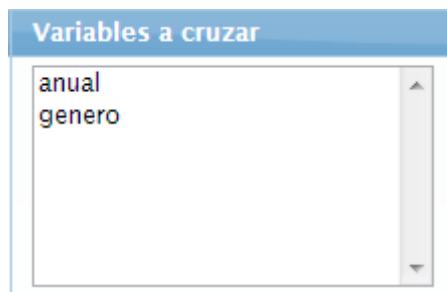


Figura F14. Variables a cruzar de un indicador

**Agregar variable a cruzar:** Se permite al usuario agregar variables al cruce de dos formas (véase figura F15):

- Dar doble clic sobre una variable del listado de ‘variables disponibles’, de esta manera la variable disponible se agrega al listado de ‘variables del cruce’ y se quita del listado de ‘variables disponibles’.
- Seleccionar una o varias variables del listado de ‘variables disponibles’ con lo que se le activa el botón → (Agregar), presionando este botón la o las variables seleccionadas de la lista de ‘variables disponibles’ se agregan al listado de ‘variables del cruce’ y se quitan del listado de ‘variables disponibles’.



Figura F15. Adición de variables disponibles a variables a cruzar

**Quitar variable a cruzar:** Para quitar una de las variables que procesara el indicador se puede realizar de dos maneras (véase figura F16):

- Dar doble clic sobre una variable del listado de ‘variables del cruce’, de esta manera la variable disponible se agrega al listado de ‘variables disponibles’ y se quita del listado de ‘variables del cruce’.
- Seleccionar una o varias variables del listado de ‘variables del cruce’ con lo que se le activa el botón ← (Quitar), presionando este botón la o las variables seleccionadas de la lista de ‘variables del cruce’ se agregan al listado de ‘variables del disponibles’ y se quitan del listado de ‘variables del cruce’.

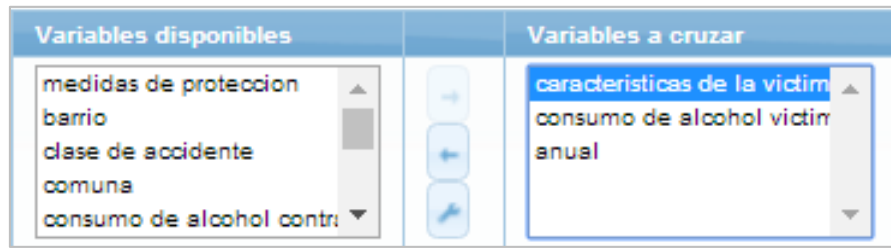


Figura F16. Quitar variables a cruzar

**Configurar variable a cruzar:** Para que se active el botón de configuraciones se debe inicialmente haber seleccionado una variable del listado de ‘variables a cruzar’. Todas las variables a cruzar pueden ser configuradas para brindar al usuario la posibilidad de seleccionar que valores de la variable desea mostrar así como configurar los rangos de tiempo en la variable hora y los rangos de edades en la variable edad. Al presionar el botón configurar se nos muestra una ventana para la configuración, los posibles casos se muestran a continuación.

- **Configuración variables categóricas:** Las variables categóricas son aquellas que contienen un conjunto de categorías, para este ejemplo tomaremos la variable ‘características de la víctima’ del tipo de evento delictivo fatal ‘muertes por accidente de tránsito’ (véase figura F17).



Figura F17. Configuración de variable categórica

La opción Quitar categoría permite eliminar categorías que el usuario no desee utilizar en el proceso de cruce de variables. Para realizar la selección de los valores a quitar podemos seleccionar una o varias categorías de la lista y hacer uso del botón ‘Quitar’, para seleccionar varios elementos de esta lista podemos hacer uso de la tecla ‘Ctrl’.

La opción Reiniciar categorías restaura todas las categorías asociadas a la variable seleccionada que hayan sido eliminadas por el usuario.

- **Configuración variable hora y variable edad:** La configuración de la variable hora y variable edad es similar a la configuración de una variable categórica en cuanto a ‘Quitar’ y ‘Reiniciar’ categorías (véase figura F18). Adicional a estas opciones se agrega una sección que nos permite insertar una nueva categoría con valores personalizados (véase figura F19).



Figura F18. Configuración variables hora (derecha) y edad (izquierda).

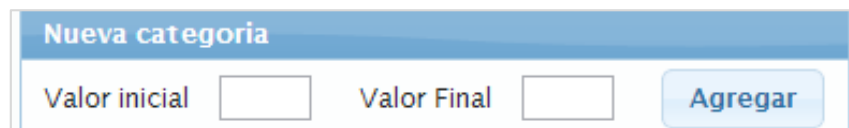


Figura F19. Sección para una nueva categoría personalizada.

Para la creación de una nueva categoría en una variable ‘hora’ o ‘edad’ se solicitan un ‘valor inicial’ y un ‘valor final’ teniendo en cuenta las siguientes consideraciones:

- Se deben insertar tanto el valor inicial como el valor final.
- El valor final debe ser mayor al final
- Los valores para hora van desde 0 hasta 23
- Los valores para edad van desde 0 hasta un número mayor que cero elegido por el usuario.
- En el caso de la variable edad se acepta como valor final la letra n, la cual nos indica que se toma el valor máximo existente.
- Cuando agregamos una nueva categoría se valida que el rango especificado no se encuentre dentro de otro ya existente (véase figuras F20 y F21).

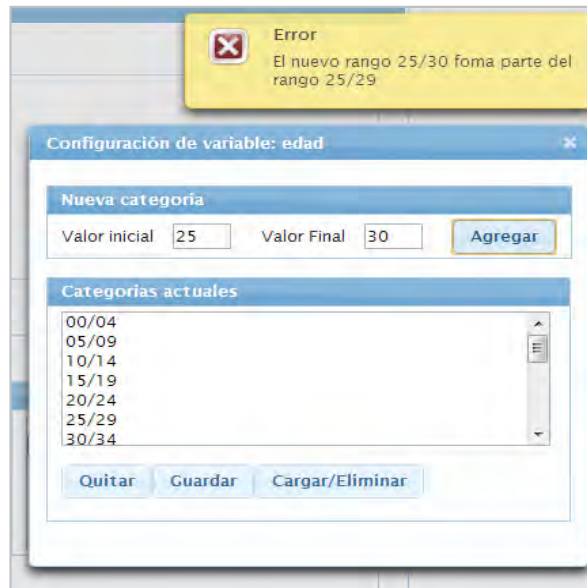


Figura F20. Nueva categoría está contenida en otra categoría existente (edad)

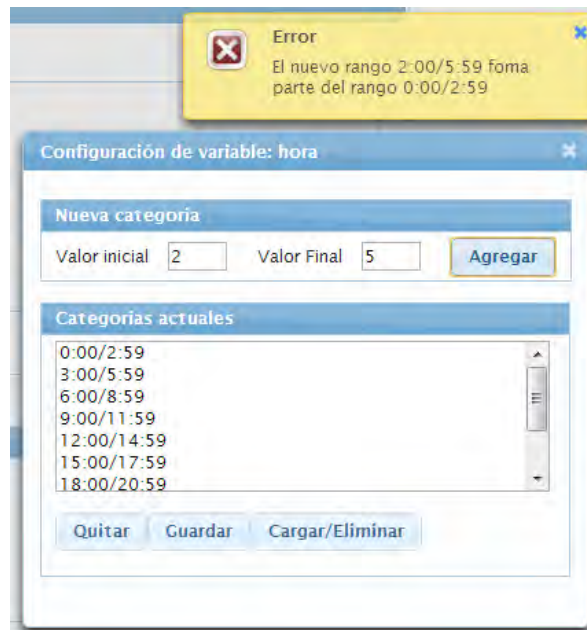


Figura F21. Nueva categoría está contenida en otra categoría existente (hora)

**Botón Procesar:** Mediante el botón '**Procesar**' se da inicio al procesamiento de los datos que serán visualizados en los mapas temáticos de puntos y de calor. Si la configuración de fechas y variables no fue realizada correctamente, el sistema informa donde el usuario debe corregir. De lo contrario, si todo está correctamente configurado, el sistema procesa los datos requeridos y despliega los

eventos delictivos que cumplan con dicha configuración en un mapa temático de puntos (véase figura F22).

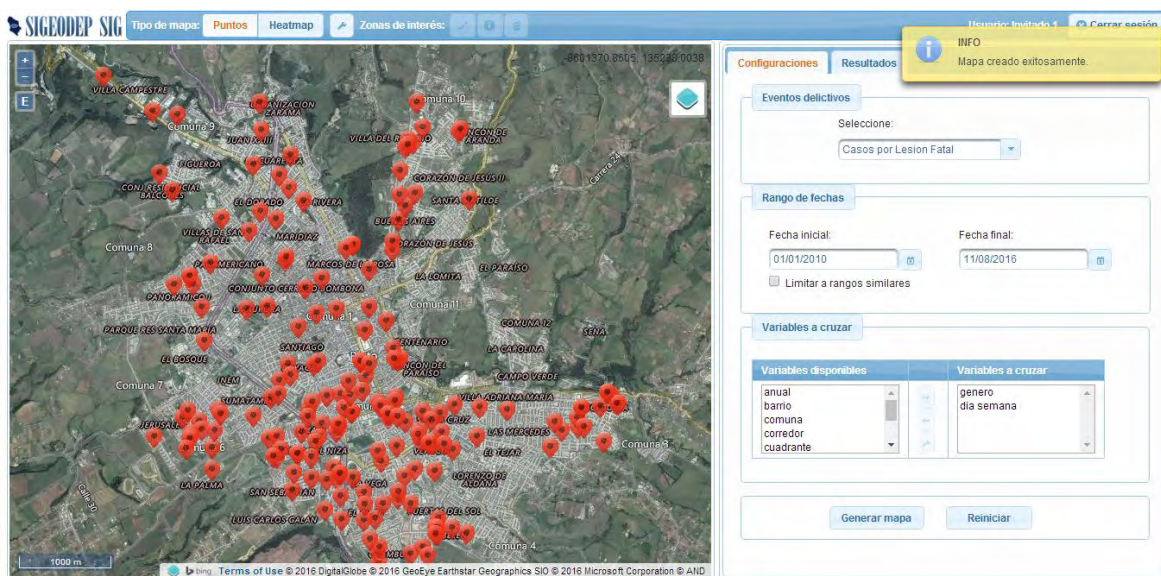


Figura F22. Mapa de puntos creado mediante cruce de variables.

**Botón Reiniciar:** De forma específica, el botón ‘Reiniciar’ realiza las siguientes operaciones:

- Reiniciar los valores de Rango de fechas y Variables a cruzar
- Reiniciar los valores de las categorías de cada variable disponible. Esto para cuando se ha configurado las categorías de una variable.
- Elimina la última capa de puntos que haya sido creada.

**Mínimo y máximo número de variables del cruce:** Por cada categoría de eventos delictivos se permite como máximo el cruce de tres (3) variables. Se deben considerar los siguientes casos para el cruce de variables:

- Las categorías ‘Casos por Lesión Fatal’ y ‘Casos por Lesión no fatal’ asignan internamente la variable ‘Tipo de lesión’, por lo tanto se permiten agregar dos (2) variables adicionales como máximo para ser cruzadas.
- En las demás categorías o Tipos de eventos delictivos Fatales y No fatales se debe agregar como mínimo una variable y como máximo tres variables.

**2.2.2. Barra de interacción con mapas temáticos:** En esta barra se ubican opciones que permiten interactuar con los mapas temáticos generados con el cruce de variables (véase figura F23).



Figura F23. Barra de interacción con mapas temáticos.

**2.2.3. Tipo de mapa:** Esta opción permite seleccionar el tipo de mapa que desea visualizarse a partir de los eventos delictivos georreferenciados a nivel de direcciones urbanas.

**Mapa de puntos:** Este es el mapa predeterminado que permite visualizar la ubicación de cada evento delictivo a nivel de direcciones urbanas. Cada punto creado corresponde a un evento delictivo, se pueden observar características particulares de la ubicación urbana de los mismos dando clic izquierdo sobre cada uno de ellos.

Entre las características mencionadas se muestran los valores que tiene el delito con respecto a las variables seleccionadas para la creación del mapa de puntos (véase figura F24).

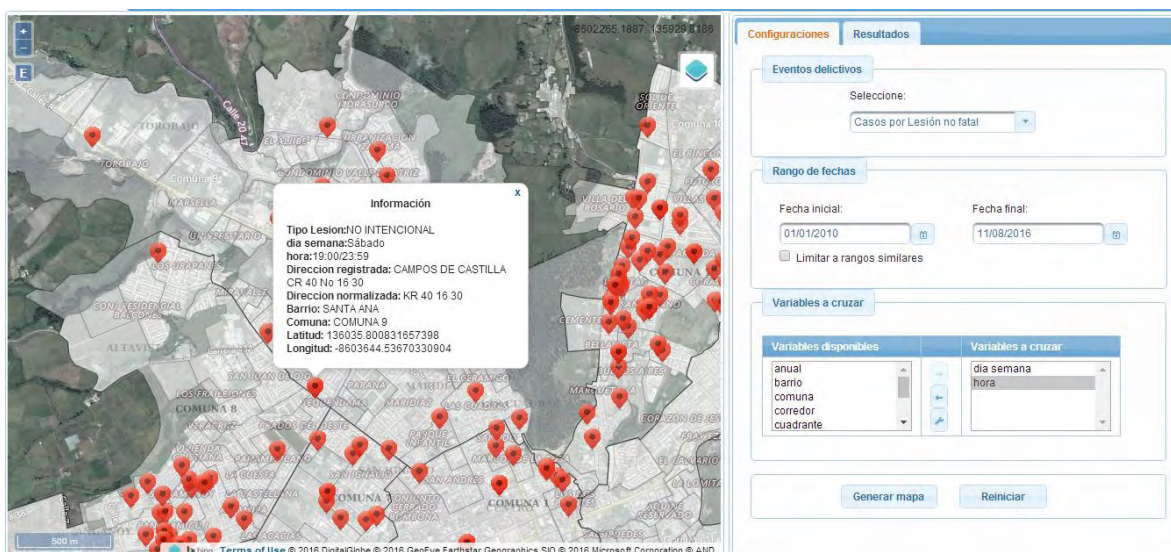


Figura F24. Visualización de eventos delictivos en mapa de puntos y sus características.

**Mapa de calor:** Este tipo de mapa agrupa los puntos generados de acuerdo a su ubicación espacial permitiendo visualizar e identificar zonas de interés o riesgo donde se presenta una gran cantidad de estos (véase figura F25), de acuerdo a una escala de colores que se describe más adelante. La visualización de estos patrones o zonas de riesgo depende directamente de tres elementos:

- La cantidad de eventos delictivos presentes en la base de datos que cumplen con las condiciones y restricciones establecidas en el cruce de variables y que fueron exitosamente procesados por el geocodificador de direcciones urbanas. Son muchos los eventos que cumplen con la configuración de variables definida por el usuario, pero su campo dirección urbana contiene valores que no pueden ser procesados por el geocodificador.
- La configuración de la opción Radio de los puntos (véase más adelante en el apartado ‘Configuración de mapa’), el mapa puede ajustarse para identificar zonas de riesgo de acuerdo a la cantidad de estos.

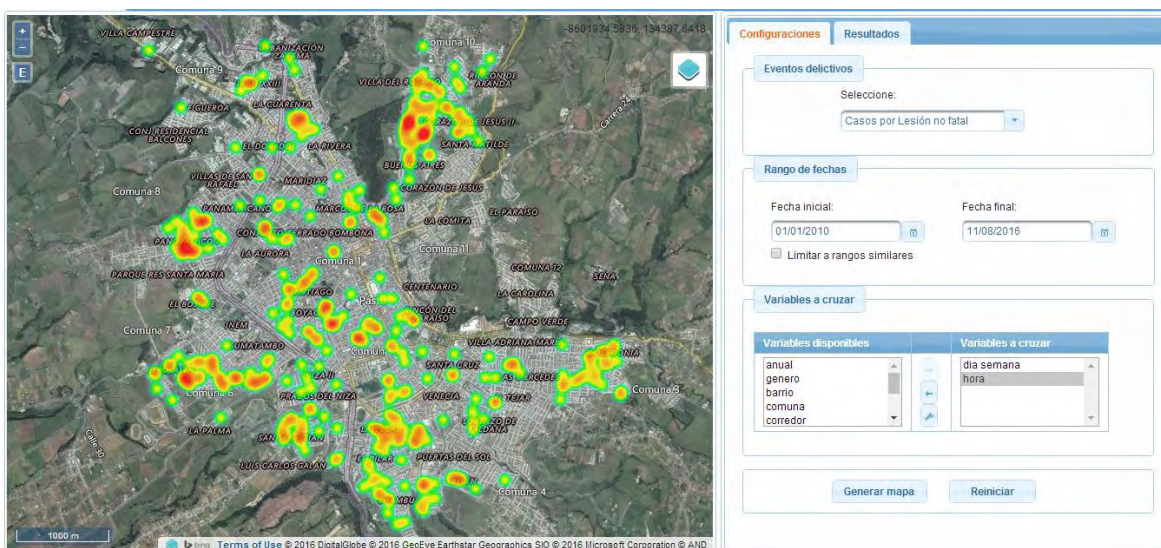


Figura F25. Mapa de calor creado a partir del cruce de variables

**Escala de colores:** Con ayuda de los mapas de calor creados por la aplicación, se puede identificar el grado de ocurrencia aproximado de eventos delictivos en zonas específicas del municipio de Pasto, esta se puede identificar de acuerdo a la distribución de colores presentada en la figura F26.

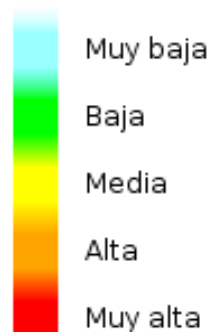



Figura F26. Escala de ocurrencia de eventos delictivos según la distribución de colores visualizada en el mapa de calor.

**Configuración de mapa:** La opción (  'configuración de mapa') muestra opciones de configuración de los mapas de puntos y de calor creados con respecto a la opacidad de visualización de los mismos (véase figura F27).

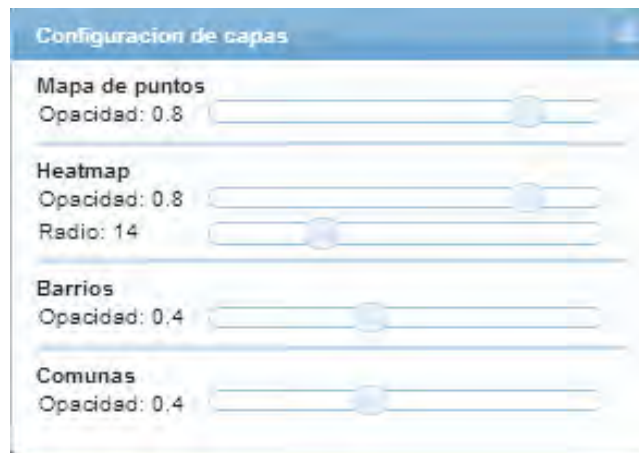


Figura F27. Configuración de opacidad de los mapas disponibles.

En el caso de los mapas de calor se tiene una opción adicional denominada 'Radio', la cual configura un atributo necesario para la creación de las zonas de riesgo. Este puede moverse para visualizar zonas de riesgo con pocos datos georreferenciados.

**Zonas de interés:** Las zonas de riesgo (de color rojo) que se pueden identificar a través de los mapas de calor, pueden ser seleccionadas para verificar la cantidad de eventos delictivos presentes mediante un conteo estadístico asociado al cruce de variables configurado por el usuario. Este

conteo puede ser realizado mediante tres opciones dispuestas por el visor cartográfico: ‘Dibujar zonas de interés’, ‘Seleccionar zona de interés’ y ‘Eliminar zonas dibujadas’. El acceso a estas opciones estará habilitado solo cuando el usuario seleccione el mapa de calor para la visualización de eventos delictivos.

**Dibujar zona de interés:** Al hacer clic sobre esta opción, el usuario accede a dibujar áreas rectangulares sobre zonas de interés con el fin de ver la cantidad de eventos delictivos ubicados. Al hacerlo el botón cambiara a color naranja, indicando que la opción ‘Dibujar áreas’ esta habilitada. De igual forma, un circulo de color azul se ubicará sobre el puntero, indicando que la opción de dibujar zonas o áreas rectangulares está habilitada (véase figura F28).

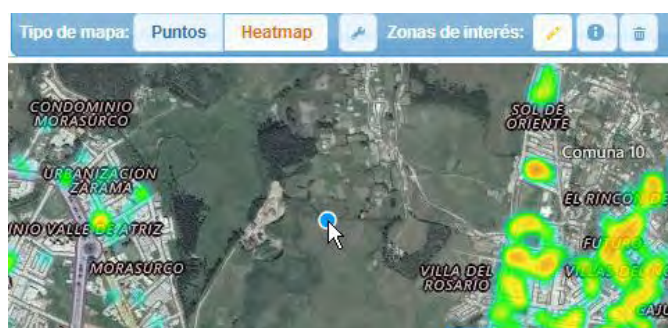


Figura F28. Forma del cursor con la opción ‘Dibujar Áreas’ habilitada.

Para dibujar estas áreas rectangulares se debe hacer un solo clic sobre el mapa el cual definirá el primer vértice del área. Al desplazar el cursor se podrá observar la forma de la figura rectangular que se está dibujando. Para terminar de dibujar el área de nuevo se hace un único clic sobre el mapa. Esta acción puede repetirse según las áreas de interés que el usuario desee consultar (véase figura F29). Para desactivar esta opción basta con hacer de nuevo clic sobre la opción ‘Dibujar áreas de interés’ o acceder a las siguientes opciones.

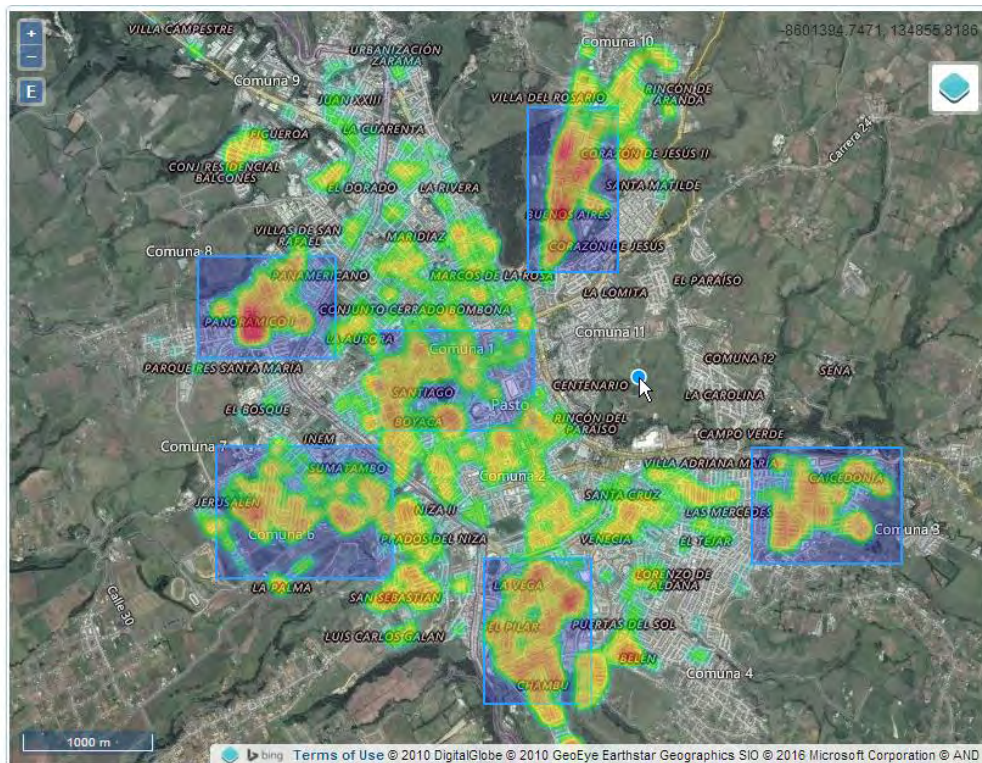


Figura F29. Ejemplos de áreas dibujadas. Opción dibujar zonas de interés está habilitada.

**Seleccionar zona de interés:** Una vez dibujadas las áreas de interés, estas pueden ser seleccionadas con la opción ‘Seleccionar zona de interés’ para visualizar el conteo de eventos delictivos presentes en esa área. Al activar esta opción, el usuario puede seleccionar cualquiera de las áreas, a partir de lo cual el aplicativo automáticamente creará un gráfico de barras básico en el panel ‘Resultados’. El grafico será creado de acuerdo a la configuración definida en la selección y cruce de variables.

La figura F30 muestra un ejemplo de selección de áreas de interés y creación de su respectivo grafico de barras.

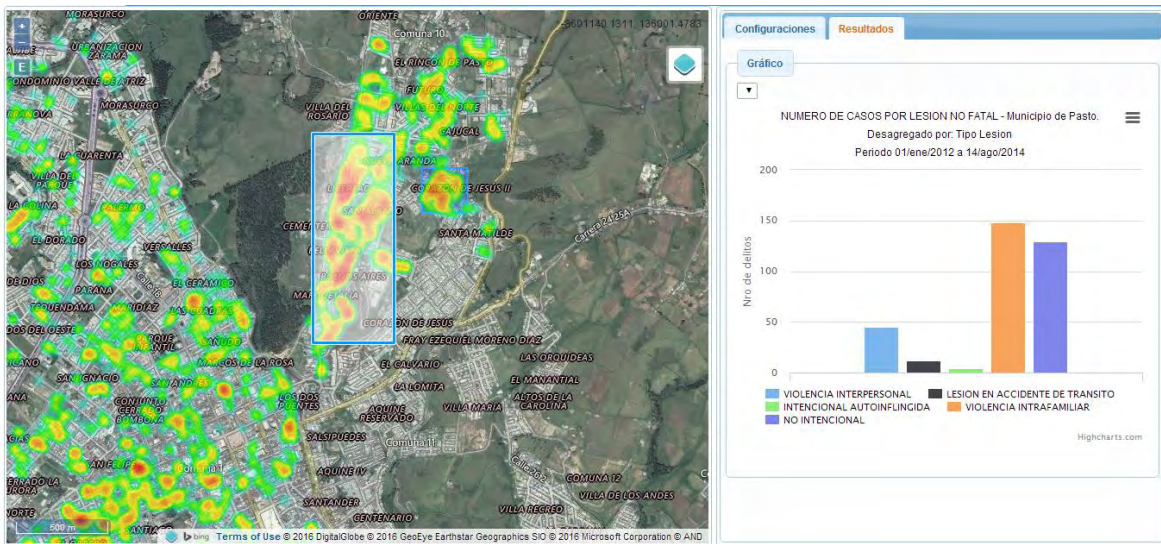


Figura F30. Resultado al activar la opción ‘Seleccionar áreas’ y seleccionar alguna de las áreas dibujadas.

**Gráfico estadístico:** La creación del gráfico de barras depende directamente del número de variables que han sido seleccionadas por el usuario. A continuación se describen algunos aspectos que deben considerarse en la creación de este gráfico:

- Como se mencionó anteriormente, las categorías de eventos delictivos ‘Casos de Lesiones de Causa Externa Fatales’ y ‘Lesiones de Causa Externa No Fatales’ manejan internamente la variable ‘Tipo de Lesión’. Esta variable será siempre ubicada en el eje horizontal.
- Si el número de variables seleccionadas es solo una (1), el grafico de barras creado será de **tipo básico**. La figura F31 muestra un ejemplo de este caso, en donde la categoría de evento delictivo seleccionada es ‘Casos de Lesiones de Causa Externa No fatales’, el rango de fechas configurado está entre las fechas 01/01/2012 y 14/08/2014 y además no fueron seleccionadas variables. El aplicativo utiliza los valores de esta variable para realizar el conteo de eventos delictivos y crear el grafico. La información relacionada con la configuración de categoría, fecha y variables se muestra en el rotulo principal del gráfico.

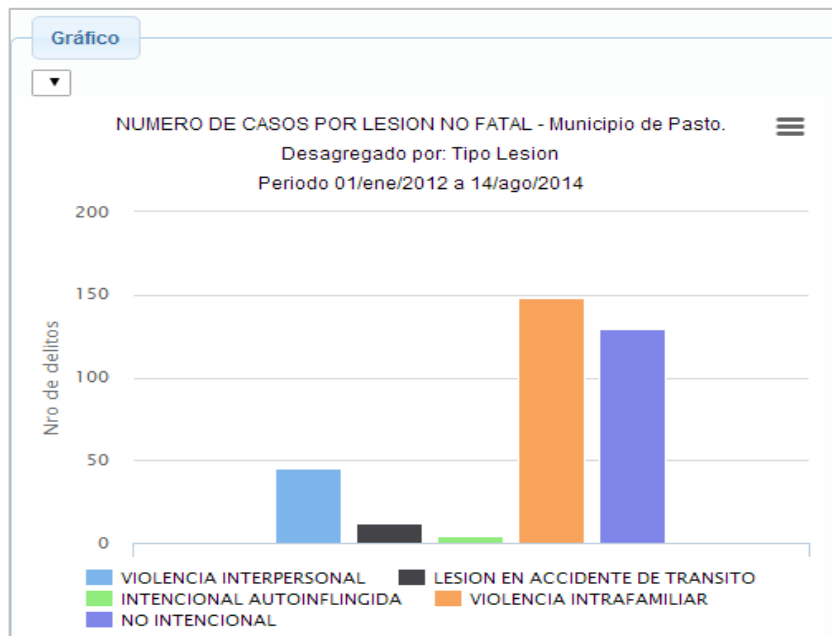


Figura F31. Gráfico de barras creado a partir de la selección de una variable.

- Si el número de variables es dos (2), el gráfico de barras generado será de **tipo agrupado**. Consideremos el anterior ejemplo, al cual se ha agregado la variable 'Género'. El aplicativo tomará los valores de la primera variable 'Tipo de Lesión', los agrupará de acuerdo a los valores de la última variable ingresada 'Genero' y realizará el conteo de acuerdo a esta agrupación. La figura F32 muestra el gráfico de barras que resultará con esta configuración.

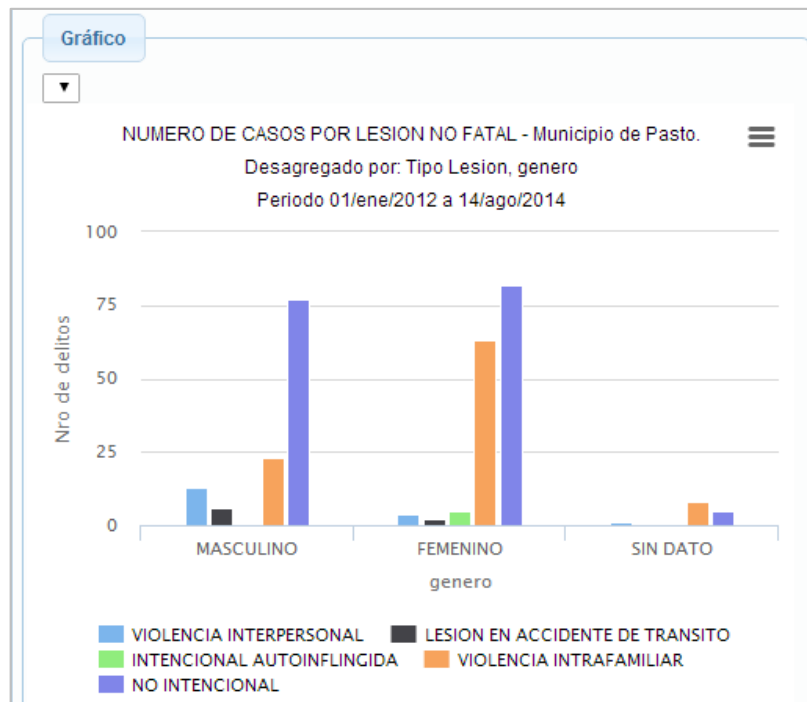


Figura F32. Gráfico de barras con agrupación, generado a partir del cruce de dos (2) variables

- Si el número de variables es tres (3), el grafico de barras generado será de **tipo agrupado con filtro**. Al agregar una tercera variable 'Año' al anterior ejemplo, generar el mapa y consultar un área de interés específica, se tendrá el mismo grafico de barras agrupado y un filtro ubicado en la parte superior del gráfico. Este filtro tomará los valores de la última variable seleccionada (en este caso 'Año o Anual') los cuales servirán de filtro para consultar la información agrupada según el filtro seleccionado. En la figura F33 se muestra el ejemplo de esta configuración, el cual muestra el conteo de eventos delictivos desagregado o agrupado por las variables 'Tipo de Lesión', 'Genero' y el año de ocurrencia filtrado al 2013.

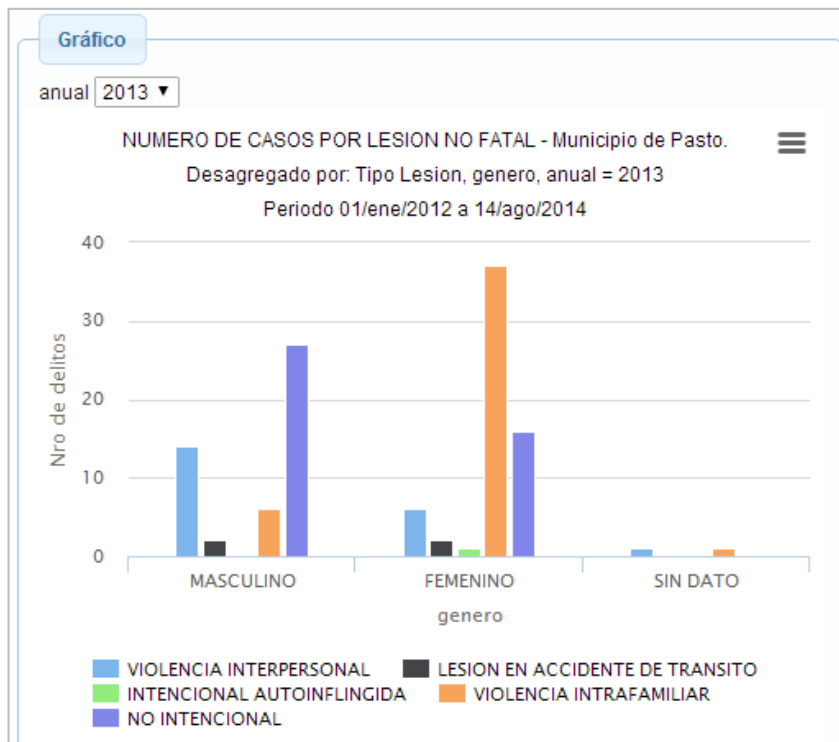


Figura F33. Gráfico de barras con agrupación y filtro (anual - parte superior) generado a partir del cruce de dos (3) variables.

Por lo anteriormente descrito, es importante considerar el orden y la cantidad permitida de variables de las variables que se cruzarán, ya que de esta forma se define el orden en el que se distribuirán los datos en el gráfico de barras al seleccionar un área de interés en el mapa de calor.

**Eliminar zonas de interés:** Al seleccionar esta opción todas las áreas creadas por el usuario serán eliminadas, permitiendo iniciar de nuevo el trazo de nuevas áreas de interés. Cabe mencionar que esta eliminación de figuras se realiza de igual forma al intercambiar el tipo de mapa.

**2.3. Archivos CSV:** A continuación se explica la funcionalidad subida y procesamiento de archivos CSV. Su implementación en SIGEODEP SIG puede observarse en la figura F34.

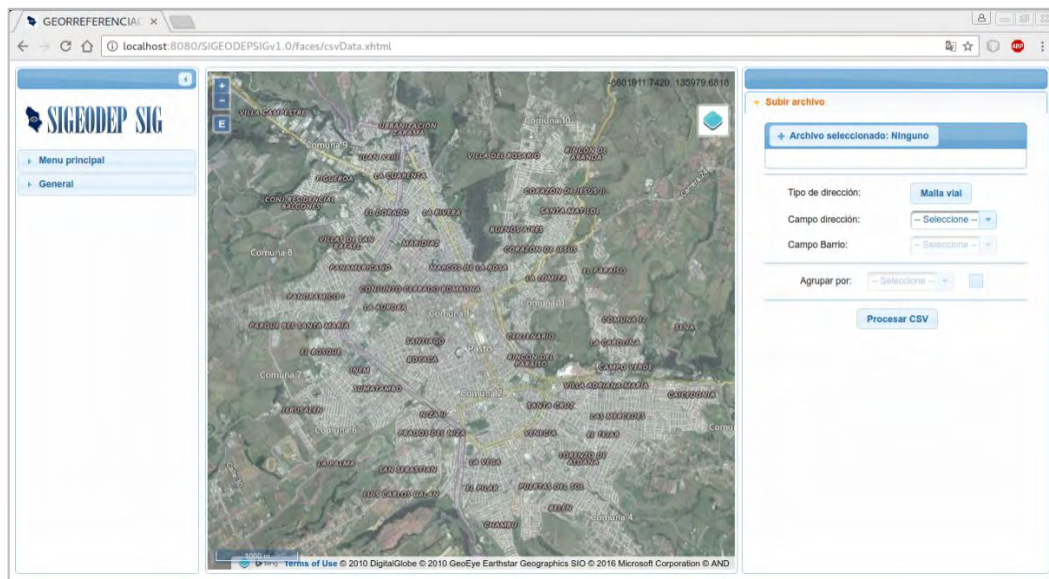


Figura F34. Procesamiento de datos CSV, pagina web implementada en SIGEODEP SIG

El panel Archivos CSV se compone de 3 opciones principales:

- **Subir archivo CSV:** En esta opción se selecciona el archivo que será procesado por el visor cartográfico. Las condiciones que debe cumplir este archivo son:
  - ✓ Los datos almacenados en este archivo deben utilizar el carácter coma (,) como delimitador.
  - ✓ Debe tener un registro cabecera (header) que identifique los nombres de cada campo.
  - ✓ Este debe tener la extensión .CSV, de lo contrario se mostrara un mensaje de archivo inválido.

El archivo será subido al servidor de la aplicación correctamente si en la etiqueta Archivo subido se muestra el nombre del archivo seleccionado.. La figura F35 muestra un ejemplo de esta subida de archivos. De igual forma se cargará las listas de selección de los campos restantes.



Figura F35. Subida correcta e incorrecta de archivos CSV.

- Selección de campos georreferenciados:** Esta opción permite seleccionar los campos del archivo CSV donde se encuentran los datos georreferenciados. En SIGEODEP SIG Se tiene disponible el tipo de campo Direcciones, aquí se especifica los campos tipo texto donde estén almacenados direcciones urbanas, según sea el caso, Malla vial y nomenclatura barrio - manzana – predio (Manzanas), que estén configuradas dentro de los datos CSV. Un ejemplo de esta configuración puede observarse en la figura F36.



Figura F36. Configuración del archivo ‘lesiones\_v2.0.csv’, el cual contiene direcciones asociadas al caso Malla vial.

- Agrupar por:** Esta opción por defecto aparecerá desactivada por lo cual los datos procesados se visualizarán georreferenciados a nivel de direcciones urbanas. La figura F37 muestra el procesamiento del archivo ‘clusters-fatales.csv’, ubicando cada registro procesado en su correspondiente ubicación.

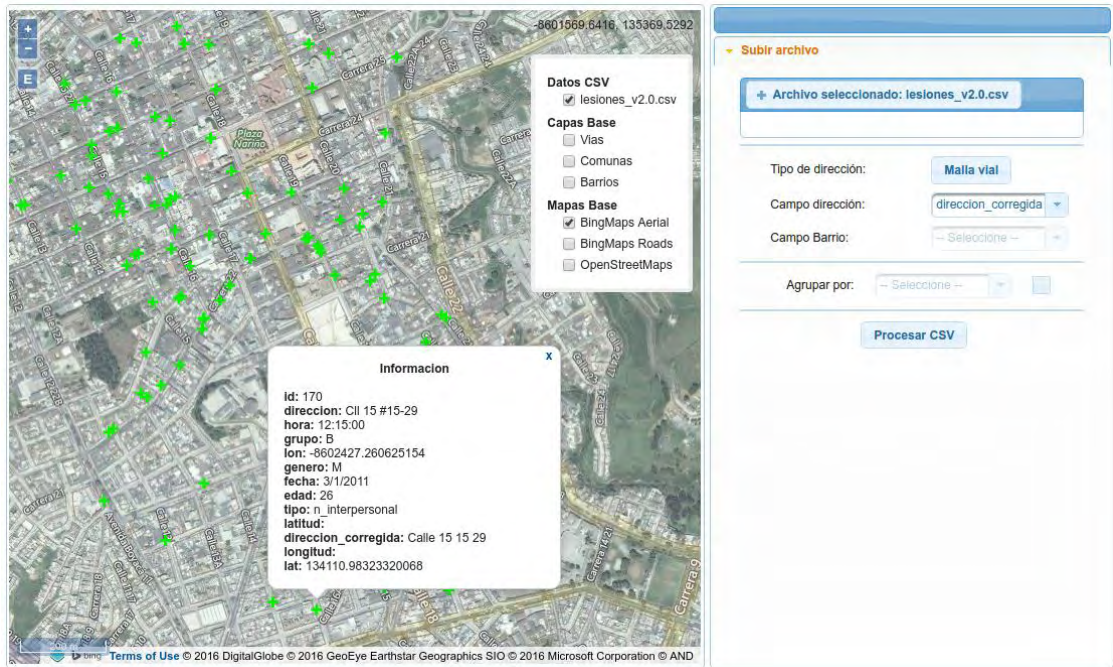


Figura F37. Procesamiento del archivo ‘clusters-fatales.csv’ sin agrupar.

Si esta opción se habilita, el visor cartográfico creará capas o mapas de visualización de acuerdo al campo seleccionado. Estas capas agruparán los datos según la característica o campo seleccionado, permitiendo deshabilitar una o varias de estas capas. La figura F38 muestra un ejemplo de esta configuración, en donde se agrupan los datos por el campo ‘Cluster’ deshabilitando las capas que corresponden a los valores ‘cluster0’ y ‘cluster1’.

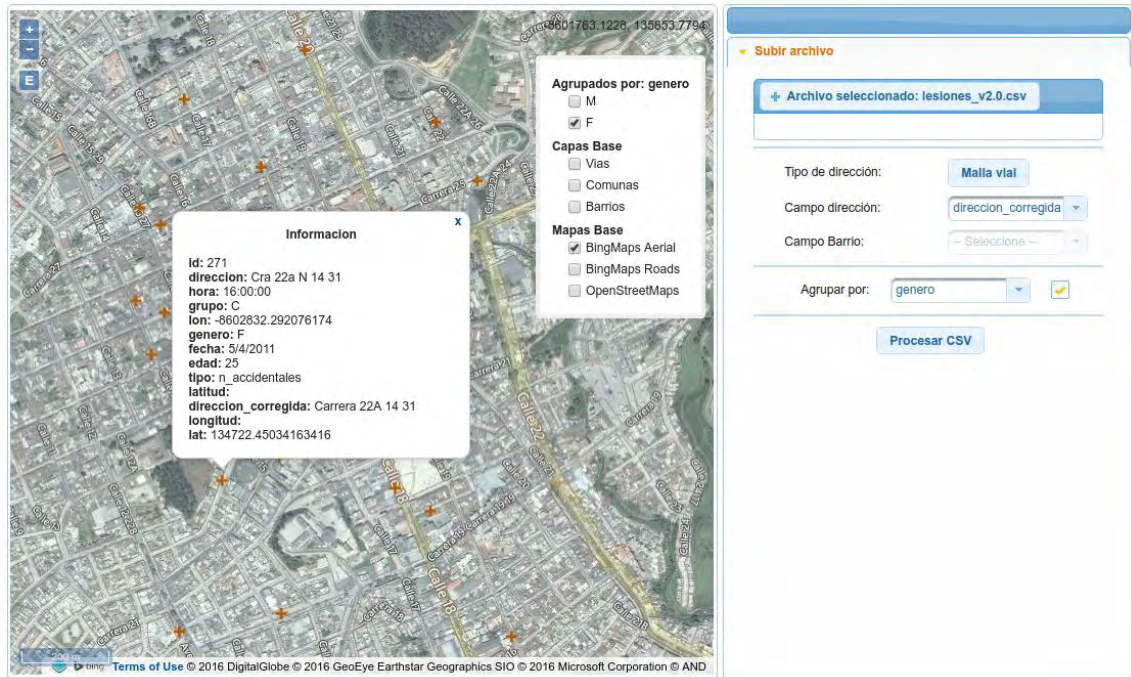


Figura F38. Procesamiento del archivo ‘lesiones\_v2.0.csv’. Los datos se agrupan según los valores del campo ‘Clusters’.

El usuario podrá visualizar a los datos asociados a cada evento o registro procesado haciendo clic en cada punto. Al hacerlo se desplegará un dialogo con cada característica asociada a dicho punto o registro.

## ANEXO G

### Análisis UML visor cartográfico SIGEODEP SIG

A continuación se describen aspectos del análisis realizado para definir las funcionalidades del visor cartográfico SIGEODEP SIG.

- 1. Actores del sistema:** Para el visor SIGEODEP SIG se definió un único tipo de usuario denominado Invitado el cual tendrá acceso sin restricciones a las funcionalidades de consulta y visualización de datos implementadas en el visor cartográfico. Este usuario podrá analizar eventos delictivos fatales y no fatales a nivel de direcciones urbanas almacenando durante el tiempo de sesión las configuraciones que haya realizado.
- 2. Funciones del sistema:** La tabla D1 muestra los requerimientos definidos para el visor cartográfico SIGEODEP SIG.

Geocodificación de eventos delictivos		
R1	Conexión con la base de datos del observatorio del delito	El sistema debe conectarse con la base de datos del observatorio del delito del municipio de Pasto para la consulta de información georreferenciada.
R2	Despliegue y selección de categoría de eventos delictivos.	El sistema debe desplegar las categorías que correspondan a los eventos delictivos fatales y no fatales, cargando previamente las variables y sus posibles valores asociadas a cada categoría.
R2.1	Cruce de variables.	El sistema debe permitir al usuario configurar variables para realizar el cruce de las mismas con el fin de analizar tendencias en los datos.
R2.2	Configuración rango de fechas.	El sistema debe permitir al usuario seleccionar un rango de fechas que los datos deben cumplir para ser procesados en el cruce de variables.
R2.3	Despliegue de variables relacionadas con la categoría de eventos delictivos seleccionada.	El sistema debe mostrar al usuario las variables asociadas a una categoría de eventos delictivos seleccionada. Cada

		variable debe tener cargados sus posibles valores
R2.3.1	Selección de variables para cruzar	El sistema debe permitir al usuario seleccionar como máximo 3 variables para ser cruzadas. Las categorías conteo general de los eventos delictivos fatales y no fatales tendrán una variable predeterminada que corresponderá al tipo de lesión.
R2.3.2	Configuración de variables seleccionadas	El sistema debe permitir al usuario configurar los valores asociados a cada variable seleccionada. El usuario puede reiniciar los valores predeterminados de cada variable.
R2.4	Procesamiento de cruce de variables	El sistema debe procesar los valores configurados por cada variable seleccionadas generando un cruce todos contra todos entre los valores y realizando el conteo de cada caso generado.
R2.4.1	Carga de eventos delictivos geocodificados de acuerdo al cruce de variables configurado.	El sistema debe almacenar los registros arrojados en el proceso de cruce de variables, los registros generados deben referenciar a las tablas de eventos delictivos geocodificados permitiendo acceder a las coordenadas geográficas de los delitos que fue posible geocodificar.
R2.5	Visualización de eventos delictivos asociados al cruce de variables configurado.	El sistema debe desplegar los registros generados con el cruce de variables a partir de las coordenadas geográficas asociadas a los eventos delictivos que fue posible geocodificar.
R2.5.1	Visualización de mapa de puntos	El sistema debe desplegar los registros generados con el cruce de variables en mapas de puntos. Al dar click sobre cada punto creado, el usuario visualizará datos asociados al delito seleccionado.
R2.5.2	Visualización de mapa de calor	El sistema debe desplegar los registros generados con el cruce de variables en mapas de calor permitiendo identificar patrones visuales en zonas donde haya una alta ocurrencia de delitos.
R2.6.1	Selección áreas de interés	El sistema debe permitir al usuario seleccionar zonas de interés mediante la creación de figuras geométricas rectangulares.

R2.6.2	Consulta de áreas de interés seleccionadas	El sistema debe permitir al usuario seleccionar las figuras geométricas creadas, procesando los eventos delictivos que se encuentren dentro de esa zona seleccionada.
R2.6.3	Generación de grafico estadístico	El sistema debe generar un gráfico estadístico a partir del conteo de eventos delictivos ubicados dentro del área seleccionada. El grafico debe desplegar información relacionada a la configuración de las variables realizada por el usuario.
R3	Carga de cartográfica correspondiente al municipio de Pasto.	El sistema debe cargar el mapa satelital del municipio de Pasto que permita visualizar las calles y vías con aproximación a la realidad.
R4	Navegación sobre los mapas cargados.	El sistema debe permitir al usuario navegar sobre los mapas temáticos cargados (ampliar, reducir la visibilidad del mapa, activación/desactivación de capas),
R5	Configuración de mapas base y capas de datos.	El sistema debe permitir al usuario configurar la visibilidad de las capas mediante un control deslizante. En el caso de los mapas de calor se adiciona un control deslizante adicional que configura el parámetro radio utilizado para la generación de mapas de calor.
<b>Archivos CSV</b>		
R6	Procesar archivos CSV	El sistema debe permitir procesar archivos separados por comas CSV con el fin de georreferenciar los registros que se encuentren almacenados.
R7	Seleccionar campos georreferenciados	El sistema debe permitir seleccionar el tipo de campos georreferenciados, los cuales pueden ser direcciones urbanas y/o coordenadas geográficas.
R7.1	Procesamiento de campo dirección urbana y barrio	El sistema debe permitir identificar el tipo de dirección urbana (Malla vial o Manzanas) presente entre los datos CSV y seleccionar los campos relacionados con direcciones urbanas, estos son dirección urbana y barrio.
R8	Agrupar datos	El sistema debe permitir agrupar los datos procesados de acuerdo a la selección de un determinado campo de datos.

R8.1	Visualizar capas de datos agrupados	El sistema debe permitir crear capas o mapas de visualización de datos geocodificados, de acuerdo a la configuración de agrupación de datos que se haya seleccionado.
Usuarios		
R9	Manejo de sesiones	El sistema debe gestionar sesiones de usuarios controlando aspectos como tiempo de inactividad o asignación de sesiones a nuevos usuarios.
R9.1	Almacenamiento temporal de configuraciones.	El sistema debe almacenar temporalmente las configuraciones realizadas por cada usuario, eliminando los datos generados y configuraciones una vez se haya terminado la sesión.

Tabla D1. Requerimientos funcionales del visor SIGEODEP SIG

- 3. Casos de uso:** Un caso de uso es una secuencia de interacciones que se desarrollarán entre un sistema y sus actores en respuesta a un evento que inicia un actor principal sobre el propio sistema.

Los casos de uso definidos que describen el funcionamiento de SIGEODEP SIG son los siguientes:

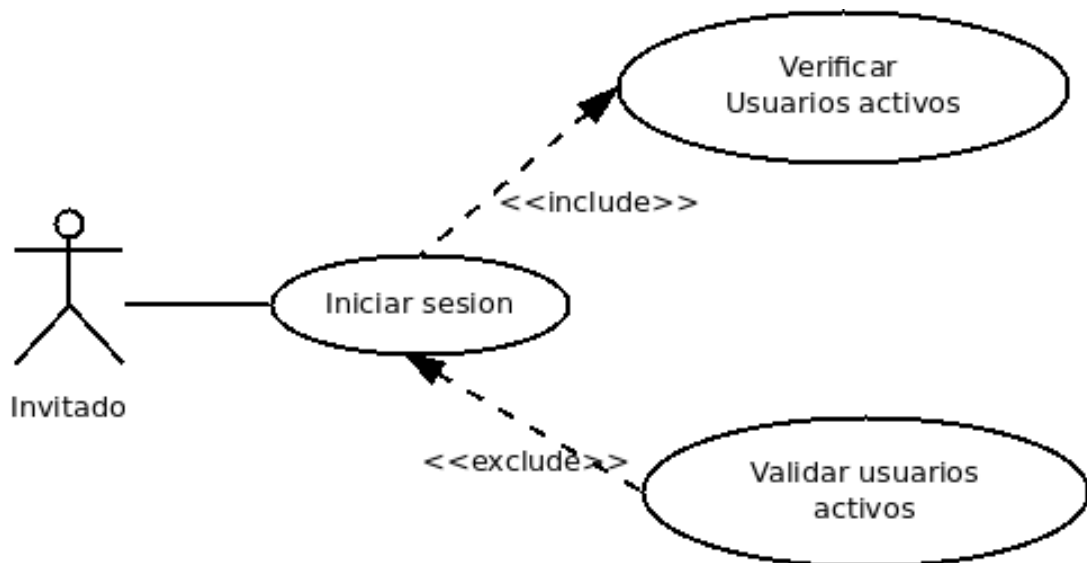
- Ingreso al visor cartográfico
  - Generación mapa de puntos mediante cruce de variables.
  - Generación de gráficos estadísticos a partir de selección de áreas de interés.
  - Procesamiento de archivos separados por comas CSV.
- Caso de Uso 1: Ingreso al visor cartográfico

Caso de uso:	Ingreso al visor cartográfico
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> </ul>
Post-condiciones:	Ingreso correcto al sistema.
Flujo Básico	

1. El sistema muestra la opción de inicio de sesión como usuario invitado
2. El usuario solicita ingresar al sistema haciendo clic en el botón inicio de sesión.
3. El sistema verifica las sesiones existentes.
4. El sistema asigna un identificador de sesión al usuario invitado.
5. El sistema agrega el nuevo usuario a la lista de usuarios conectados.
6. El sistema dirige al usuario al menú de inicio.

La figura G1 muestra el diagrama de caso de uso.

Figura G1. Diagrama de caso de uso Ingreso al visor cartográfico



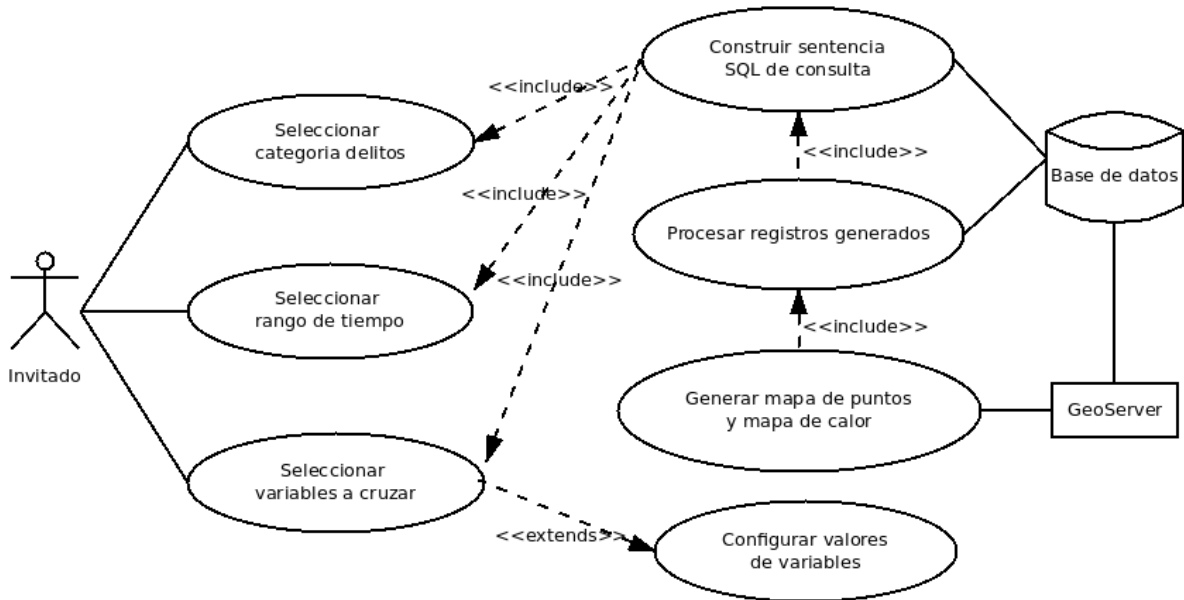
- Caso de Uso 2: Generación mapa temáticos mediante cruce de variables.

Caso de uso:	Generación de mapas temáticos mediante cruce de variables
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Eventos delictivos geocodificados a nivel de direcciones urbanas</li> <li>▪ Extensión PostGIS instalada en la base de datos.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> </ul>
Post-condiciones:	Generación correcta del mapa de puntos asociado al cruce de variables configurado por el usuario.
<b>Flujo Básico</b>	
<ol style="list-style-type: none"> <li>1. En el menú principal, el Usuario selecciona la opción Análisis de Lesiones de Causa Externa.</li> <li>2. El Sistema despliega la interfaz del visor cartográfico con el mapa satelital del municipio de Pasto y el panel de configuración de variables.</li> <li>3. El Sistema lista las categorías disponibles por cada tipo de lesión de causa externa.</li> <li>4. El Sistema lista el conjunto de variables seleccionables de categoría.</li> <li>5. El Usuario selecciona la categoría Lesiones de Causa Externa Fatales o la categoría Lesiones de Causa Externa No fatales.</li> <li>6. El usuario ingresa el periodo de tiempo que será analizado.</li> <li>7. El usuario selecciona una primera variable.</li> <li>8. El sistema lista de nuevo el conjunto de variables seleccionables sin incluir la escogida previamente.</li> <li>9. El Usuario selecciona una segunda variable.</li> <li>10. El Usuario selecciona la opción de generar mapa.</li> <li>11. El Sistema consulta la base de datos, genera registros por cada combinación existente de los valores de las variables.</li> <li>12. El Sistema identifica los registros que cumple con cada registro generado con el cruce de variables.</li> <li>13. El Sistema busca y almacena los registros que fueron geocodificados.</li> <li>14. El servidor de mapas consulta los registros que fueron almacenados por el Sistema.</li> <li>15. El Servidor de Mapas crea el mapa temático de puntos de acuerdo a las variables configuradas.</li> <li>16. El Sistema despliega el mapa temático de puntos previamente solicitado al Servidor de Mapas.</li> <li>17. El Sistema muestra el mensaje Mapa creado exitosamente.</li> </ol>	
<b>Flujos alternos</b>	
5a. Si el Usuario selecciona una categoría diferente, podrá seleccionar entre 1 y 3 variables.	

<p>9a. El Usuario podrá reiniciar y restablecer las configuraciones predeterminadas.</p> <p>13a. Si Sistema no encuentra registros geocodificados, el sistema informa que hubo registros que coinciden con el cruce de variables configurado que no fue posible visualizar.</p> <p>17a. El usuario podrá intercambiar la visualización de datos a mapas de calor.</p>
<p><b>Flujos de excepción</b></p> <p>6a. El sistema validará que el periodo final no sea anterior al periodo inicial. En dicho caso se despliega un mensaje de error.</p> <p>10a. El Sistema verificara que la categoría seleccionada sea Lesiones de Causa Externa Fatales o Lesiones de Causa Externa No fatales. En dicho caso validará que haya un máximo de dos variables seleccionadas. De haberse seleccionado más variables se muestra un mensaje de error.</p> <p>10b. El Sistema verificara que la categoría seleccionada no este entre Lesiones de Causa Externa Fatales y Lesiones de Causa Externa No fatales. En dicho caso validará que haya un máximo de tres variables seleccionadas. De haberse seleccionado más variables se muestra un mensaje de error.</p> <p>10c. El Sistema verificara que la categoría seleccionada sea Lesiones de Causa Externa Fatales o Lesiones de Causa Externa No fatales. En dicho caso validará que haya mínimo una variable seleccionada. De haberse seleccionado más variables se muestra un mensaje de error.</p>

La figura G2 muestra el diagrama de caso de uso.

Figura G2. Diagrama de caso de uso: Generación de mapas temáticos mediante cruce de variables.



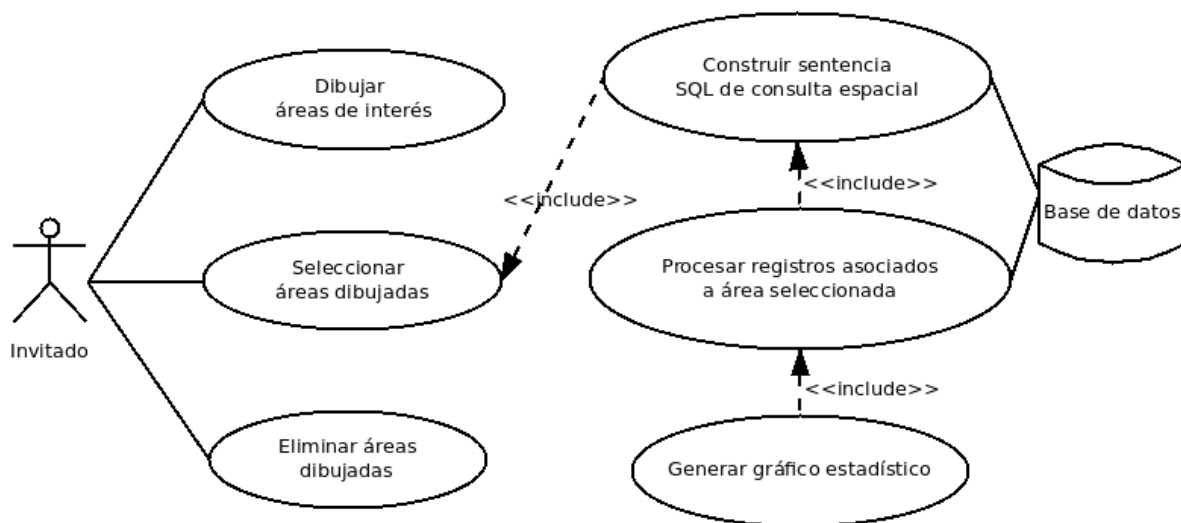
- Caso de uso 3: Generación de gráficos estadísticos a partir de selección de áreas de interés.

Caso de uso:	Generación de gráficos estadísticos a partir de selección de áreas de interés.
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> <li>▪ Proceso de cruce de variables y visualización de datos ejecutados.</li> <li>▪ Mapa de calor generado y seleccionado.</li> </ul>
Post-condiciones:	Ingreso correcto al sistema.
Flujo Básico	
<ol style="list-style-type: none"> <li>1. El Sistema muestra el mapa de calor generado a partir del previo cruce de variables.</li> <li>2. El Usuario selecciona la opción Dibujar Áreas de Interés.</li> <li>3. El Usuario traza figuras geométricas rectangulares sobre el mapa de calor en las zonas donde desea consultar y analizar datos.</li> <li>4. El Usuario selecciona la opción Seleccionar Áreas de Interés y selecciona una de las áreas dibujadas.</li> <li>5. El Sistema captura las coordenadas geográficas de los vértices de la figura geométrica seleccionada.</li> </ol>	

6. El Sistema genera y ejecuta la consulta espacial SQL que seleccionara los datos del cruce de variables asociados a los eventos delictivos situados dentro de la figura geométrica.
  7. El Sistema recupera los datos solicitados.
  8. El Sistema configura las variables requeridas para la generación del grafico estadístico considerando la configuración de variables cruzadas y sus valores realizada.
  9. El Sistema activa una pestaña de visualización de gráficos estadísticos
  10. El Sistema despliega el grafico estadístico mostrando el conteo de delitos realizado sobre el área geográfica seleccionada.
- Flujos alternos:
- 10a. El usuario puede eliminar las figuras geométricas dibujadas.

La figura G3 muestra el diagrama de caso de uso.

Figura G3. Diagrama de caso de uso Generación de gráficos estadísticos a partir de selección de áreas de interés.



- Caso de uso Procesamiento de archivos CSV.

Caso de uso:	Procesamiento de archivos CSV
Actores del sistema:	Usuario invitado
Pre-condiciones:	<ul style="list-style-type: none"> <li>▪ Conexión correcta a base de datos del observatorio del delito.</li> <li>▪ Geocodificador de direcciones urbanas instalado.</li> <li>▪ Servidor de publicación de mapas temáticos en ejecución.</li> </ul>
Post-condiciones:	Mapa(s) de puntos asociado(s) al procesamiento de los registros encontrados en el archivo CSV.
Flujo Básico	
<ol style="list-style-type: none"> <li>1. El sistema despliega el formulario de selección y procesamiento de archivos CSV.</li> <li>2. El usuario selecciona la opción ‘Seleccionar archivo’ y busca el archivo que desea procesar.</li> <li>3. El sistema procesa el archivo seleccionado cargando los nombres de cada campo encontrado en las listas de selección de campos.</li> <li>4. El usuario configura las listas de campos de acuerdo a la estructura del archivo CSV y presiona el botón ‘Procesar’.</li> <li>5. El sistema procesa cada registro de acuerdo a los campos georreferenciados seleccionados.</li> <li>6. El sistema despliega la capa de puntos con cada evento o registro procesado informando a usuario la cantidad de registros procesados exitosamente.</li> </ol>	
Flujo alternativo	
<ol style="list-style-type: none"> <li>2a. El sistema verifica que el archivo no tiene la extensión .csv.</li> <li>2b. El sistema informa al usuario que el archivo no es válido.</li> <li>4a. Si la opción Agrupar datos es habilitada, el usuario debe seleccionar el campo por el cual se agruparan los datos.</li> <li>4b. El sistema procesa cada registro de acuerdo a los campos georreferenciados seleccionados.</li> <li>4c. El sistema despliega la capa de puntos con cada evento o registro procesado informando a usuario la cantidad de registros procesados exitosamente.</li> </ol>	

La figura D2 muestra el diagrama de caso de uso.

Figura D2. Diagrama de caso de uso Procesamiento de archivos CSV

