

BIG DATA PARA LA PRIORIZACIÓN DE ZONAS DE ATENCIÓN A EMERGENCIAS  
CAUSADAS POR INUNDACIONES EN BOGOTÁ - COLOMBIA: USO DE LAS REDES  
SOCIALES

Priorización de zonas afectadas por inundaciones usando datos de las redes sociales.

Danny Lilian Godoy Godoy

Laura Alejandra González Gómez

Universidad de Ciencias Aplicadas y Ambientales

16 de junio de 2019

Nota

Trabajo de investigación. Profesor: José Alejandro Salamanca García. Ingeniería Geográfica y  
Ambiental. Universidad de Ciencias Aplicadas y Ambientales.

La correspondencia relacionada con este documento deberá ser enviada a:

[dgodoy@udca.edu.co](mailto:dgodoy@udca.edu.co)/[laura.agonzalez@udca.edu.co](mailto:laura.agonzalez@udca.edu.co)

BIG DATA PARA LA PRIORIZACIÓN DE ZONAS DE ATENCIÓN A EMERGENCIAS  
CAUSADAS POR INUNDACIONES EN BOGOTÁ - COLOMBIA: USO DE LAS REDES  
SOCIALES

DANNY LILIAN GODOY GODOY  
LAURA ALEJANDRA GONZÁLEZ GÓMEZ

Trabajo de Grado

José Alejandro Salamanca García

UNIVERSIDAD DE CIENCIAS APLICADAS Y AMBIENTALES  
INGENIERÍA GEOGRÁFICA Y AMBIENTAL  
BOGOTA  
2019

## CONTENIDO

Glosario.....	6
Lista de acrónimos .....	8
RESUMEN .....	10
1. INTRODUCCIÓN.....	11
1.1 Descripción del área de estudio.....	13
2. PLANTEAMIENTO DEL PROBLEMA .....	16
2.1 Problema de investigación .....	16
2.2 Hipótesis.....	18
2.3 Objetivo general y objetivos específicos.....	19
2.3.1 Objetivo general .....	19
2.3.2 Objetivos específicos .....	19
2.4 Justificación.....	20
3. REVISIÓN DE LITERATURA. ....	23
3.1 Antecedentes .....	23
3.2 Marco conceptual .....	39
3.2.1 Big Data.....	39
3.2.2 Redes sociales.....	41
3.2.3 Inundaciones .....	41
3.2.4 Desastre .....	42
3.2.5 Respuesta .....	42
3.3 Estado del arte o del conocimiento .....	43
4. METODOLOGÍA.....	46
4.1 Fase 1: revisión de metodologías y procedimientos.....	46
4.2 Fase 2: análisis DOFA.....	53
4.3 Fase 3: generación de guía metodológica .....	55
5. RESULTADOS Y ANÁLISIS DE RESULTADOS .....	57
5.1 Comparación de metodologías de aplicación del Big Data en otros países .....	57
5.2 Análisis DOFA de la subdirección de manejo de emergencias y desastres del IDIGER.	
67	
5.2.1 ESTRATEGIAS F.O .....	81
5.2.2 ESTRATEGIAS D.O.....	82
5.2.3 ESTRATEGIAS F.A .....	83

5.2.4 ESTRATEGIAS D.A.....	84
5.3 Guía metodológica para la priorización de zonas afectadas por inundaciones .....	84
6. DISCUSIÓN .....	110
CONCLUSIONES .....	119
RECOMENDACIONES.....	122
ANEXOS .....	123
Anexo 1. Entorno de trabajo de Eclipse.....	123
Anexo 2. Lenguaje para definir variables y parámetros de conexión .....	123
Anexo 3. Pestaña de preprocesamiento de Weka.....	124
Anexo 4. Partes del archivo ARFF.....	124
Anexo 5. Fuente de los datos requeridos en la priorización.....	125
Anexo 6. Análisis multicriterio para la instalación de estaciones meteorológicas: red de apoyo y monitoreo.....	126

## LISTADO DE FIGURAS

Figura 1. Mapa de ubicación de la zona de estudio .....	15
Figura 2. Inundaciones Fluviales .....	29
Figura 3. Inundaciones Urbanas .....	29
Figura 4. Inundaciones Pluviales .....	30
Figura 5. Zonificación hidrográfica de Bogotá.....	35
Figura 6. Zonificación de amenaza por inundación en Bogotá .....	38
Figura 7. Marco metodológico.....	46
Figura 8. Big Data Analytics para la respuesta ante desastres y recuperación a través del análisis de sentimientos.....	50
Figura 9. Metodología 1: Inundación a través del lente de la actividad del teléfono móvil....	59
Figura 10. Metodología 2: Explorando el uso de las redes sociales durante la inundación de 2014 en Malasia .....	60
Figura 11. Metodología 3: Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres .....	61
Figura 12. Metodología 4: Dominar las redes sociales: un análisis de las comunicaciones del condado de Jefferson durante las inundaciones de Colorado 2013 .....	62
Figura 13. Metodología 5: Chatter on The Red: ¿Qué revela la amenaza sobre la vida social de la información generada en microblogs? .....	63

Figura 14. Selector GUI de Weka.....	87
Figura 15. Zonas de intersección en áreas de amenaza .....	93
Figura 16. Modelo de elevación digital para la ciudad de Bogotá.....	95
Figura 17. Modelo de elevación digital con zonas de amenaza por inundación.....	96
Figura 18. Creación de polígonos de Thiessen desde entidades de puntos .....	98
Figura 20. Porcentaje de viviendas afectadas por tipo de inundación. ....	100
Figura 19. Aplicación polígonos de Thiessen.....	101
Figura 21. Aplicación interpolación IDW .....	102
Figura 22. Zonas de amenaza por inundación en Bogotá. ....	103
Figura 23. Zonas de amenaza por inundación en el barrio el Tintal-Bogotá. ....	104
Figura 24. Aplicación de los polígonos de Thiessen. ....	105
Figura 25. Aplicación de interpolación IDW.....	106
Figura 26. Pasos metodológicos para priorizar zonas de atención frente a inundaciones mediante el uso de redes sociales.....	108
Figura 27. Diagrama de flujo general .....	109

#### LISTADO DE TABLAS

Tabla 1. Análisis ONI para escenarios de inundaciones.....	31
Tabla 2. Cifras de la emergencia invernal 2010-2011 .....	32
Tabla 3. Comparación: modelos metodológicos de aplicación Big Data. ....	64
Tabla 4. Matriz DOFA de la subdirección de manejo de emergencias y desastres del IDIGER. ....	69
Tabla 5. Área de intersección en zonas de amenaza. ....	92

## Glosario

- **Amenaza:** peligro latente de que un evento de origen natural o antrópico ocurra con la severidad necesaria para causar impactos, daños y pérdidas en la vida de las personas, infraestructura, bienes, prestación de servicios y recursos naturales. (Ley 1523 de 2012).
- **Análisis de sentimientos:** es la aplicación de técnicas de analítica y del procesamiento del lenguaje natural con el fin de extraer información subjetiva de diversas fuentes como texto, documentos y otras. Este análisis es muy utilizado actualmente en medios sociales como blogs, wikis, redes sociales y demás para deducir comportamientos de clientes, hábitos de compra o sistemas de recomendación. (Joyanes, 2013).
- **Big Data:** hace parte de la nueva generación de tecnologías y arquitecturas diseñadas con el fin de darle un valor económico a los grandes volúmenes de datos a partir de su captura y análisis a una gran velocidad. (IDC, 2012).
- **Datos estructurados:** datos con formato fijo que poseen campos fijos. Son datos que se pueden encontrar en las bases de datos relacionales, hojas de cálculo y archivos y se utilizan en las operaciones diarias de las organizaciones. (Joyanes, 2013).
- **Datos no estructurados:** son datos que no tienen campos fijos como por ejemplo los archivos de audio, vídeo, fotografías o formatos de texto libre tales como correos electrónicos, SMS, artículos, libros, mensajería instantánea y otros. (Joyanes, 2013).
- **Datos semiestructurados:** datos que no tienen formatos fijos, pero contienen etiquetas y marcadores que permiten separar datos elementos. Un ejemplo de estos datos son el texto de etiquetas de XML y HTML. (Joyanes, 2013).
- **Emergencia:** es una situación que se caracteriza por la alteración de las condiciones normales de una comunidad, se origina por un evento adverso y requiere de una reacción inmediata por parte de las instituciones del Estado, medios de comunicación y de toda la comunidad. (Ley 1523 de 2012).

- **Internet de las cosas:** red de objetos que conectados entre sí se envían datos mutuamente a través de protocolos de Internet IP. Es decir que los objetos se comunican entre sí porque a cada uno de ellos se le atribuye una dirección de internet IP. (Joyanes, 2013)
- **Inundación:** es un evento en donde la precipitación constante hace que el nivel normal del cauce natural o artificial de un cuerpo de agua aumente de tal forma que invade otros sitios, ocasionando daños a la población, infraestructura, recursos ambientales y demás. (CENAPRED, 2019)
- **Machine Learning:** también conocido como aprendizaje automático, es la ciencia de hacer que las computadoras detecten patrones y hagan conexiones mediante un conjunto de datos, a través de la aplicación de algoritmos, con el fin de que la máquina tenga sus propios parámetros para desarrollar una tarea y resolver problemas por sí sola. (UNI, 2019).
- **Minería de datos:** es una de las técnicas más utilizadas en el análisis de datos y se está adaptando al análisis de grandes datos. Esta incluye dos grandes categorías que influyen en la web y en la documentación y la biblioteconomía: minería web y minería de textos. Estas técnicas ayudan al análisis eficiente de grandes volúmenes de datos y actualmente se considera una nueva categoría de intercambio de datos entre sensores conocida como máquina a máquina (M2M). (Joyanes, 2013)
- **Respuesta:** realización de las actividades necesarias para atender una emergencia como accesibilidad, transporte, telecomunicaciones, evaluación de daños y de necesidades, salud, búsqueda y rescate, albergues, alimentación, servicios públicos y demás. (Ley 1523 de 2012)

## Lista de acrónimos

API. Application Programming Interface (Interfaz de programación de aplicaciones)

CDR. Call Detail Records (Registro de detalles de llamadas)

CMC. Comunicación Mediada por Computadora

DANE. Departamento Administrativo Nacional de Estadística

DEM. Digital Elevation Models (Modelo Digital de Elevaciones)

DOFA. Análisis de Debilidades, Oportunidades, Fortalezas y Amenazas.

GRD. Gestión del Riesgo de Desastres

IDEAM. Instituto de Hidrología, Meteorología y Estudios Ambientales

IDECA. Infraestructura de Datos Espaciales para el Distrito Capital

IDW. Interpolación de la Distancia Inversa Ponderada

IDE. Infraestructura de Datos Espaciales

IGAC. Instituto Geográfico Agustín Codazzi

TI. Tecnología de la información

NUSE. Número Único de Seguridad y Emergencias

OMM. Organización Meteorológica Mundial

SIG. Sistema de Información Geográfica

SIGOT. Sistema de Información Geográfica para la Planeación y el Ordenamiento Territorial

TIC. Tecnologías de la información y las comunicaciones

TIN. Triangular irregular networks (red irregular de triángulos)

USGS. Servicio Geológico de los Estados Unidos

WMS. Web Map Service

## RESUMEN

En los procesos de respuesta ante los desastres dados por las inundaciones en Bogotá Colombia; la utilización de la información derivada de los datos Big Data generados en las redes sociales, conocidos como datos desestructurados, en la subdirección de manejo de emergencias y desastres del Instituto Distrital de Gestión de Riesgos y Cambio Climático (IDIGER); apoyará las respuestas institucionales, entendidas como: la toma de decisiones, los tiempos de atención y la adecuada gestión de recursos por medio de la priorización de zonas afectadas, mediante la recolección, procesamiento y análisis de la información obtenida de forma rápida y eficaz.

La siguiente investigación, pretende determinar cómo el aprovechamiento de la información obtenida de los datos generados de redes sociales, es de utilidad para priorizar zonas afectadas y apoyar la toma de decisiones dentro de las respuestas institucionales del IDIGER frente a desastres generados por inundaciones, a través de la identificación de las acciones desarrolladas durante la atención a una emergencia; una revisión bibliográfica de la metodología y procedimientos dentro de los modelos de aplicación sobre el uso de la información, que compara tales acciones con los proyectos realizados en otros países, así como la proposición de una guía metodológica que plantee opciones para analizar la información obtenida, enfocándose en la determinación y representación cartográfica de zonas afectadas por desastres asociados a inundaciones; priorizando dichas zonas y apoyando las respuestas institucionales.

Lo anterior tiene como resultado: una revisión bibliográfica de las metodologías y procedimientos dentro de modelos de aplicación del uso de la información obtenida de los datos generados en las redes sociales, frente al manejo de desastres causados por inundaciones, y su respectivo análisis comparativo; un diagnóstico de las actividades realizadas dentro de la subdirección de manejo de emergencias y desastres del IDIGER para atender una emergencia por inundación, y la formulación de una propuesta que consta de una guía metodológica enfocada en la priorización de zonas afectadas por inundaciones, que hace uso de herramientas cartográficas, así como de métodos de extracción de la información directamente de los datos generados en las redes sociales y el análisis de dicha información para apoyar la toma de decisiones en las respuestas institucionales.

## 1. INTRODUCCIÓN.

El Big Data hace parte de una nueva generación de tecnologías en la que la recolección, almacenamiento, procesamiento y análisis de grandes volúmenes de datos a alta velocidad, permite una mejor toma de decisiones en las respuestas y la generación de valor agregado a la calidad y análisis de la información de cualquier institución o empresa.

Actualmente existe un mayor acceso a las nuevas tecnologías y a la información por parte de usuarios, que se caracterizan por ser personas particulares que no necesariamente son especializadas para generar, capturar, almacenar y analizar grandes cantidades de información extraída de grandes volúmenes de datos de forma automática y rápida. Esta es una tendencia tecnológica que se ha venido consolidando durante los años 2011 y 2012 con el uso de Big Data en las organizaciones, empresas y en general en la sociedad, con la aparición de las redes sociales, aumento de la banda ancha, geolocalización, Internet de las cosas y la computación en la nube. (Aguilar, 2013).

Todos los días en el mundo se generan grandes cantidades de información personal con llamadas telefónicas, las páginas de internet que visitan dichos usuarios, es decir, la huella digital, las fotos, vídeos, estados, intereses que se comparten en Facebook o Twitter, los mensajes de WhatsApp, los millones de correos que se envían y se reciben, localizaciones y muchas otras aplicaciones que hacen referencia a los datos en la web y de los medios sociales. Puede sumarse la información dada por la transferencia de datos entre los objetos que se comunican entre sí por sensores (Internet de las cosas) o la mayoría de instituciones en el mundo que tienen todos sus datos digitalizados, actualizados y en línea.

Lo anterior representa lo que es la aplicación del Big Data en general y aunque existen diferentes definiciones por la forma en que los sectores pueden aprovechar las características del Big Data como: el volumen, velocidad, variedad, veracidad y el valor de los datos, se sabe que este es un concepto nuevo en constante crecimiento que transforma la infraestructura de las tecnologías de la información tradicionales.

En el campo geográfico el Big Data puede ser utilizado desde un enfoque social, donde las personas puedan participar aportando y generando información en tiempo real, desde sus dispositivos móviles, sobre las situaciones de emergencia que están viviendo y deben ser

atendidas de forma rápida y eficaz. Es decir, que el aprovechamiento de la información obtenida de los datos generados en las redes sociales, con Big Data permite determinar la cantidad de población afectada, el comportamiento social durante un desastre, así como la priorización de zonas afectadas y el apoyo en la toma de decisiones dentro de las respuestas institucionales.

En Colombia pueden existir diversos sectores que están teniendo en cuenta las ventajas que brinda el Big Data, pero al ser un tema reciente y novedoso, hay un bajo conocimiento de sus características, funcionamiento, y aplicabilidad. Es por esto que el siguiente trabajo de investigación visualiza el Big Data como una oportunidad de aplicación en materia de atención de desastres por inundaciones, enfocado en la priorización de zonas afectadas, mediante el uso y representación cartográfica de la información obtenida de los datos generados en las redes sociales, entendiéndose como datos desestructurados<sup>1</sup>, que permitan apoyar los procesos de toma de decisiones, contribuyendo en la mejora dentro de las respuestas institucionales frente a desastres causados por inundaciones.

Cabe aclarar que según la Ley 1523 de 2012 el manejo de desastres es uno de los procesos de la gestión de riesgo que está compuesto por la preparación para la respuesta a emergencias, la ejecución de dicha respuesta, y la preparación y ejecución de la recuperación. En este caso; este trabajo solo se enfocará en la ejecución de la respuesta, la cual está definida por la misma ley, y consta de respuestas institucionales como: la ejecución de las actividades necesarias para la atención de la emergencia como accesibilidad y transporte, telecomunicaciones, evaluación de daños y análisis de necesidades, salud y saneamiento básico, búsqueda y rescate, extinción de incendios y manejo de materiales peligrosos, albergues y alimentación, servicios públicos, seguridad y convivencia, aspectos financieros y legales, información pública y el manejo general de la respuesta, entre otros.

Este trabajo de investigación tiene como objetivo proponer el uso del BIG DATA, a partir de los datos generados en las redes sociales, para la priorización de zonas de atención frente a

---

<sup>1</sup> Son datos que se almacenan como objetos sin una estructura uniforme y con poco o ningún control sobre ellos. Los datos de texto, fotografía, audio, video, cartas, imágenes digitales, mensajes de texto y otros. En esta fuente, el 80% de la información de las organizaciones se encuentra completamente esparcida

inundaciones, como herramienta de gestión para la subdirección de manejo de emergencias y desastres del IDIGER.; Lo anterior tendrá como resultados: una revisión bibliográfica de las metodologías y procedimientos dentro de modelos de aplicación del uso de la información obtenida de los datos generados en las redes sociales, frente al manejo de desastres causados por inundaciones, y su respectivo análisis comparativo; un diagnóstico de las actividades realizadas dentro de la subdirección de manejo de emergencias y desastres del IDIGER para atender una emergencia por inundación; y la formulación de una propuesta que consta de una guía metodológica enfocada en la priorización de zonas afectadas por inundaciones, que hace uso de herramientas cartográficas, así como de métodos de extracción de la información directamente de los datos generados en las redes sociales y el análisis de dicha información para mejorar la toma de decisiones en las respuestas institucionales.

### 1.1 Descripción del área de estudio

Para efectos prácticos de este trabajo de investigación, se escogió como área de estudio el barrio el Tintal, ya que la muestra representativa de los puntos de las viviendas se encuentra sobre las 3 zonas de influencia de amenaza por inundación que son: alta, media y baja.

Este barrio también conocido como ciudad Tintal o nueva castilla es una ciudadela ubicada en la localidad de Fontibón al occidente de la ciudad de Bogotá<sup>2</sup> capital del departamento de Cundinamarca<sup>3</sup> y distrito capital de Colombia (ver figura 1), es uno de los barrios limítrofes entre las localidades de Fontibón (9) y Kennedy (8). Se caracteriza por tener un desarrollo residencial dirigido a ofrecer soluciones de vivienda de interés social. En otras palabras, el barrio el Tintal cuenta con una elevada dinámica constructiva habitacional en propiedad horizontal. (Galindo, 2013), a pesar de esto, el Tintal presenta un alto riesgo por encharcamiento debido al taponamiento de las redes de alcantarillado, originado por el vertimiento de residuos sólidos ordinarios.

---

<sup>2</sup> Capital de la república de Colombia y también capital del departamento de Cundinamarca. Se encuentra configurada como distrito capital.

<sup>3</sup> El departamento de Cundinamarca es uno de los 32 Departamentos que conforman la república de Colombia. Se encuentra ubicado en el centro del país, en la región Andina; al norte limita con Boyacá, al este con Casanare, al sur con Meta y Huila y al oeste con Tolima y Caldas.

Además, al ser un barrio limítrofe entre dos localidades, presenta un mayor riesgo por inundaciones. Ejemplos de lo anterior, fue lo ocurrido en uno de los sectores más cercanos al barrio el Tintal, conocido como Tierra Buena, donde se presentaron graves afectaciones por el desbordamiento del Río Bogotá, en la segunda temporada de lluvias del año 2011 (CLGR-CC, 2018). Así mismo, para esa misma época, en diciembre de 2011, en el sector de la estación elevadora de Fontibón, hubo un desbordamiento del río Fucha, lo que ocasionó la devuelta de aguas del sistema de drenaje y la inundación de las partes bajas de Kennedy y Bosa. (IDIGER, 2019).

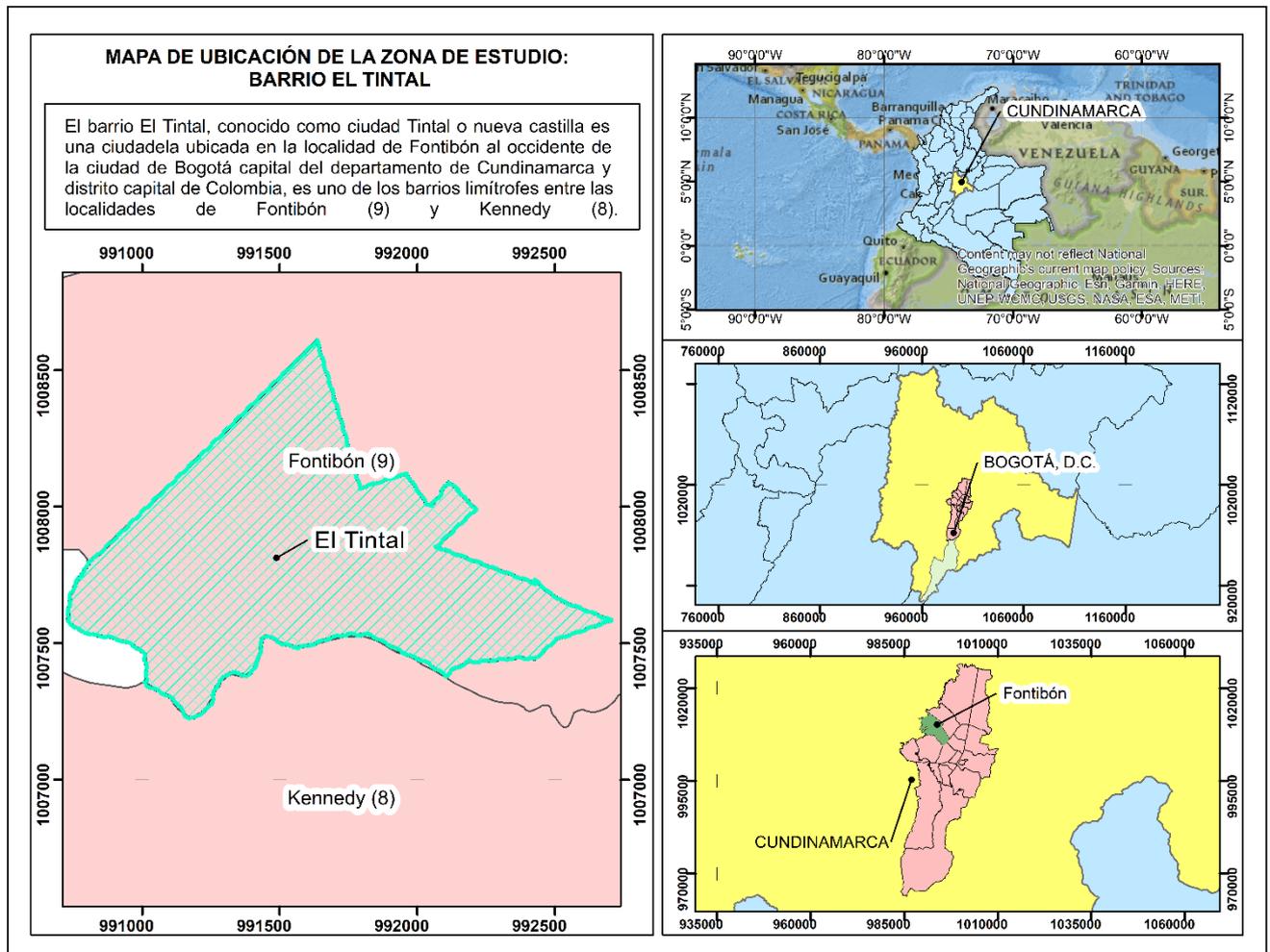
Lo anterior, ocasiona riesgos y afectaciones a la salud ya que no toda el agua que provoca las inundaciones provienen necesariamente de las lluvias o por desbordamientos de ríos no contaminados, dicha agua por lo general se encuentra contaminada con aguas residuales, productos químicos, y basuras que representan un potencial riesgo para la salud, como lo son: objetos puntiagudos, vidrios y metales oxidados. Las aguas residuales contaminadas ocasionan grandes afectaciones en la piel de las partes del cuerpo que se encuentren sumergidas en las aguas de las zonas inundadas durante periodos prolongados de tiempo. Así mismo, el agua contaminada presente en las inundaciones es un vector potencial en la transmisión de enfermedades, en ciudades como Bogotá, son más comunes los ataques de diarrea y otros problemas de estómago relacionados con el consumo de agua y alimentos contaminados, de la misma forma el uso y manipulación de artículos sumergidos también puede causar problemas serios de estómago, especialmente en los niños.

Así mismo, la exposición continua a aguas de inundación contaminadas puede aumentar el riesgo de enfermedades en la piel como: erupciones cutáneas, problemas de oído como: otitis, nariz, garganta y conjuntivitis; según la organización mundial de la salud (OMS), ninguno de los problemas anteriormente nombrados puede verse sujeto a epidemias, no obstante, las enfermedades estomacales y respiratorias si están propensas a epidemias, si se considera como agravante el hecho de que en una situación de desastre como las inundaciones, las personas afectadas deben compartir espacios pequeños con múltiples personas.

En barrios de Bogotá ubicados en las rondas de los ríos, especialmente en ríos tan contaminados como lo es la cuenca media del río Bogotá, en donde según (Revista semana, 2015), este río presenta una contaminación que no da tregua, ya que se encuentran desde sanitarios, hasta toda clase de botellas plásticas, siendo uno de los ríos más contaminados del mundo, cuya

contaminación inicia desde los primeros tramos de su nacimiento, cuando pasa por los municipios de villapinzón y se contamina con los residuos de las curtiembres, sin embargo la mayor contaminación la recibe cuando pasa por la ciudad de Bogotá en donde desembocan los demás ríos contaminados que atraviesan Bogotá de oriente a occidente, como lo son: el rio Salitre, Fucha y Tunjuelito; las inundaciones son también un problema para la salud pública, generan pérdidas económicas y afectaciones sociales de orden público.

Figura 1. Mapa de ubicación de la zona de estudio



## 2. PLANTEAMIENTO DEL PROBLEMA

### 2.1 Problema de investigación

El desaprovechamiento de la información que se obtiene de los datos generados diariamente por todo tipo de cliente<sup>4</sup> o usuario<sup>5</sup> con acceso a internet, y más aún la información generada bajo situaciones adversas, es una de las grandes desventajas a la que se enfrentan las entidades, empresas y demás, ya que el reto es saber aprovechar el análisis de dicha información y hacer de la misma una ventaja tanto competitiva como comparativa para las mismas. En la actualidad el procesamiento y aprovechamiento de datos masivos adquiere un rol importante si se enfoca hacia el progreso de la sociedad. El análisis de este tipo de información puede influir en la toma de decisiones dentro de las respuestas institucionales, permitiendo que la recolección de información sea canalizada hacia una mejor respuesta ante eventos catastróficos (UPM, 2017), causados específicamente por inundaciones.

Se debe tener en cuenta que desde los años 90's, la mayoría de las grandes ciudades de países de América Latina; entre ellos Colombia: Bogotá, se encuentran llevando a cabo proyectos enfocados en el mejoramiento de infraestructura básica, para hacerle frente a la insuficiencia de obras encaminadas hacia la mejora y ampliación de las redes de alcantarillado existentes, y por consiguiente para el tratamiento de las inundaciones, sin embargo la manera en la que actualmente se abordan las problemáticas del drenaje pluvial, junto con la gestión y atención de las inundaciones en Colombia, muestra aun un grado de retraso. (Bertoni, 2013).

Es por esto que, para efectos prácticos dentro de esta investigación, se encaminará el aprovechamiento de la información obtenida de los datos Big Data generados en las redes sociales para la respuesta frente a inundaciones, ya que la vulnerabilidad física, económica y sociocultural ante dichos eventos ha ido incrementando gradualmente, especialmente en zonas urbanas y periurbanas del país, como consecuencia de la urbanización desmedida, la

---

<sup>4</sup> Es para quien se van a crear o diseñar los productos o servicios de una organización, como lo puede ser una empresa. (Ramirez, 2018)

<sup>5</sup> El **cliente** es el sujeto que adquiere o paga por un producto o servicio. Así mismo, se lo concibe como una persona que disfruta con cierta regularidad de ese servicio o producto, razón por la que se habla de una fidelización con la marca. (Ramirez, 2018)

ineficiencia en el alcantarillado y los numerosos asentamientos ilegales en zonas de riesgo (Idarraga, 2010).

Es decir, que la respuesta a emergencias presenta una baja planificación, teniendo en cuenta que esta varía dependiendo del tipo de desastre, los programas y recursos disponibles. En Colombia hay diversidad de eventos naturales peligrosos que pueden ser la causa de desastres, los cuales pueden diferenciarse según los tipos de fenómenos y su severidad. Por consiguiente, la magnitud de los daños, también cambia en cada región por los diferentes factores de vulnerabilidad tanto físicos, sociales, económicos, políticos y culturales. De igual forma, en el momento de ocurrencia de algún evento, la respuesta del Estado está sujeta a las disposiciones fiscales existentes, esto condiciona que las estrategias de atención con programas y recursos sean específicos. (Banco Mundial Colombia, 2012, pág. 258).

Lo anterior, sumado a una falta de claridad en las políticas sobre la responsabilidad del gobierno y de los agentes frente a las respuestas a emergencias, representa poco avance en: la identificación de escenarios de posibles pérdidas, la definición de tareas a ejecutar para el manejo de emergencias, los mecanismos de coordinación y la utilización de recursos. Lo anterior debido a que en cada evento se adoptan diferentes acciones (Banco Mundial Colombia, 2012) y existe poca integración entre la información de las entidades oficiales; reduciendo la efectividad y rapidez en los tiempos de atención.

Es allí donde la incorporación del análisis de la información obtenida de los datos Big Data juega un papel importante dentro de la gestión del riesgo de desastres, enfocado específicamente a la priorización de zonas afectadas y apoyo en la toma de decisiones en las respuestas institucionales por parte del IDIGER frente a emergencias causadas por inundaciones, ya que desde y dentro de los diversos sectores de la sociedad se está dando paso a la generación y almacenamiento de cantidades masivas de datos, los cuales reflejan el comportamiento de las personas en la cotidianidad o en situaciones adversas y pueden ser aprovechados en la priorización y atención inmediata a damnificados y/o afectados.

En consecuencia, a lo anterior, es preciso determinar la manera en la que la información, obtenida del análisis de los datos Big Data generados en las redes sociales, se puede ver y entender como una herramienta transversal que proporcione información verídica y confiable,

y permitan apoyar el proceso de toma de decisiones frente a la atención inmediata de poblaciones afectadas.

Cabe aclarar que en Colombia el tema de explotación de datos masivos conocidos como Big Data no hacía parte de las agendas gubernamentales, ni se encontraba dentro de los puntos claves en el desarrollo sectorial en los planes de nacionales de desarrollo (CONPES 3920), explicando así la baja disponibilidad de datos públicos digitales, causada al mismo tiempo por la poca digitalización de información analógica y la reducida publicación de los datos.

De lo anterior se pueden identificar diversas interrelaciones que condicionan el aprovechamiento del Big Data dentro de las instituciones públicas asociadas con la respuesta a emergencias en el manejo de inundaciones en Bogotá. En ese orden de ideas, si se tiene en cuenta la ausencia parcial de marcos legales que regulan la anonimización de la información personal de los usuarios que alimentan de manera diaria en tiempo real a los Big Data; la información como uno de los pilares fundamentales en los procesos internos de cualquier entidad, así como factor principal para el alcance de una interdisciplinariedad y transdisciplinariedad; o considerando que en Colombia, cada entidad es única y se especializa en un campo específico de análisis e investigación; consecuentemente: ¿Cómo se priorizan las zonas de atención a emergencias causadas por inundaciones en Bogotá a partir de los datos obtenidos en las redes sociales? ; ya que se encontró que faltan aún muchas herramientas y mecanismos que integren la información entre las entidades públicas y las comunidades durante la atención de situaciones de emergencia.

## 2.2 Hipótesis

A continuación, se presentan las hipótesis generadas, teniendo en cuenta que las hipótesis parten de suposiciones sobre los posibles resultados de la investigación derivadas del análisis del problema y la teoría propuesta para la tesis en cuestión, se presentaran dos hipótesis las cuales son una respuesta provisional al planteamiento del problema anteriormente descrito:

- A. Los datos generados en las redes sociales se posicionan como una de las principales fuentes de información, siendo utilizados como apoyo en la priorización de zonas afectadas, y en los procesos de toma de decisiones, contribuyendo en la mejora dentro de las respuestas institucionales frente a desastres causados por inundaciones.

- B. Una guía metodológica sobre el uso y aprovechamiento de la información obtenida de los datos generados en las redes sociales para priorizar zonas afectadas por inundaciones, sirve como un apoyo dentro de los procedimientos de la subdirección de manejo de emergencias y desastres del IDIGER con el fin de que mejore significativamente los tiempos de atención y tome decisiones efectivas en la determinación de la cantidad, pertinencia y organización de los insumos necesarios para las poblaciones damnificadas y/o afectadas; lo anterior, entendido como respuestas institucionales.

## 2.3 Objetivo general y objetivos específicos

### 2.3.1 Objetivo general

Proponer el uso del Big Data, a partir de los datos generados en las redes sociales, para la priorización de zonas de atención frente a inundaciones, como herramienta de gestión para la subdirección de manejo de emergencias y desastres del IDIGER.

### 2.3.2 Objetivos específicos

1. Revisar las metodologías y procedimientos sobre el uso de las redes sociales, frente al manejo de inundaciones en otras ciudades.
2. Analizar la estructura orgánica y subprocesos en la subdirección de manejo de emergencias y desastres del IDIGER.
3. Generar una guía metodológica para priorizar zonas de atención frente a inundaciones mediante el uso de redes sociales, en el marco de las funciones de la subdirección de manejo de emergencias y desastres del IDIGER.

## 2.4 Justificación

El siguiente estudio presenta, desde un enfoque teórico, un encuadre del aprovechamiento de la información obtenida a partir de los datos generados en las redes sociales con Big Data encaminado hacia la priorización de zonas afectadas por desastres causados por inundaciones.

Es preciso señalar que, en el mundo, en los últimos 6 años China, Estados Unidos, la Unión Europea, Australia, Corea del sur, Reino Unido, Japón y Francia tuvieron un avance en la implementación de una política pública nacional de aprovechamiento de datos, para que la adecuación de condiciones de cada país aumentara los beneficios sociales y económicos y disminuyera los riesgos relacionados con la explotación de datos. Mientras que, en Colombia, hace 20 años se vio la necesidad de manejar las tecnologías de la información y las comunicaciones, para la gestión gubernamental y el incremento de la eficiencia en los procesos; solo se reconoció la importancia del valor social y económico que generan los datos hasta el Plan Nacional de Desarrollo 2014-2018 *Todos por un nuevo país*.

Es decir que, en Colombia, pese a los avances en la provisión de normativas, para que las entidades públicas suministren los datos de su gestión e intercambien información, aún no se alcanzan altos niveles de datos digitales, el aumento de su explotación y la generación de un valor socioeconómico, porque este tipo de aprovechamiento de datos no se ha reconocido como un activo central de la economía digital. (CONPES 3920).

La importancia en el análisis de la información obtenida a partir de los datos desestructurados Big Data provenientes de las redes sociales, en general, radica en que, al verse como una fuente confiable de datos generados por las personas en tiempo real mediante la interacción de las mismas en las redes sociales, lo posicionan como una estrategia y una estructura innovadora, que permite la gestión y análisis de grandes volúmenes de datos procedentes de redes sociales a una velocidad considerable, también ayuda al aumento de la competitividad, productividad, rentabilidad, integración de información y mayor rendimiento y asertividad, haciendo más fácil la toma de decisiones frente a las respuestas institucionales.

El aumento de la competitividad de cualquier organización es una de las grandes ventajas en la utilización de la información obtenida de los datos Big Data, debido a que la adecuada obtención y análisis de las nuevas fuentes de datos contribuye al alcance de conocimientos

nuevos y al ofrecimiento de servicios con un valor añadido a los usuarios. McAfee y Brynjolfsson (2012) afirman que “las empresas y organizaciones que inyectan Big Data y analítica de Big Data en sus operaciones muestran tasas de productividad, efectividad y rentabilidad que oscilan en el orden del 5 al 6 %; más altas que aquellas de la competencia o compañías homólogas.”

Este aumento obligatoriamente necesita de una integración de la información extraída, donde se combinen datos tradicionales con datos Big Data, para que exista un enfoque holístico (Joyanes, 2013) y la información se convierta en uno de los bienes más valiosos para cualquier organización. En este caso se trabajará con la subdirección de manejo de emergencias y desastres del IDIGER, ya que es el encargado de orientar la ejecución de respuestas a desastres en Bogotá (Concejo de Bogotá, 2014).

Así mismo, esta investigación se enfocará principalmente en las inundaciones, debido a que son las más frecuentes en el país y en la ciudad (CEPAL, 2007) y sus afectaciones han generado grandes impactos económicos y sociales. (UNGRD, 2018) Por ejemplo, los daños como consecuencia del fenómeno de la Niña, en los años 2010-2011, alcanzan un valor de 6.052 millones de dólares, siendo el sector de hábitat (44%) y el sector de infraestructura (38%), los sectores con mayor daño. (CEPAL, 2012).

Lo anterior demuestra que ciudades como Bogotá presentan un alto riesgo de afectación por inundaciones dentro de las zonas urbanas y periurbanas, en donde los altos índices de urbanización, han alterado las condiciones de absorción pluvial normales del suelo, la construcción y asentamientos humanos en zonas de riesgo, y la falta de cultura ciudadana sumada a la deficiencia en el aseo urbano, que altera y tapona el funcionamiento normal del alcantarillado.

En resumen, la realización de la siguiente investigación puede tener un impacto positivo y un aporte auténtico, tanto para las entidades públicas encargadas de la atención de desastres y la sociedad, porque a través de una visión comparativa, con otros países que han tenido éxito en la aplicación del Big Data dentro del aprovechamiento e integración de la información obtenida de los datos provenientes de las redes sociales, se obtendrán propuestas de uso que servirán como una base para mejorar la eficacia en las acciones dadas en la respuesta a desastres por inundaciones. Esto indica que la información será una herramienta destinada a proporcionar

una respuesta apropiada, oportuna y ágil, a las personas perjudicadas por una emergencia; y algo fundamental en la valoración de necesidades y daños, la organización y la toma de decisiones cuando se movilizan recursos nacionales e internacionales. (UNGRD, 2018).

### 3. REVISIÓN DE LITERATURA.

#### 3.1 Antecedentes

En Colombia no existen proyectos relacionados con el aprovechamiento de la información obtenida a partir de los datos desestructurados Big Data provenientes de las redes sociales, en la atención de desastres causados por inundaciones, pero existen antecedentes en la normatividad relacionada con las tecnologías de la información.

Según el documento CONPES 3920 de la política de explotación de datos (Big Data), en Colombia los gobiernos, durante los años 90 y a comienzos del año 2000, se concentraron en aumentar la eficiencia en la gestión de las entidades públicas y privadas mediante el uso de las TIC.

Las intervenciones gubernamentales se realizaron en periodos de tiempo diferentes en los cuales no se consideraba el valor social y económico de los datos y se pueden agrupar en Gobierno Electrónico, eficiencia administrativa, gestión documental y protección de datos. Además, hay normas que hacen referencia a los reportes de datos por entidades públicas y privadas y reportes de información y su impacto en la generación y compartición de datos.

Lo anterior se resume en un marco jurídico sobre explotación de datos que regula 5 materias que son: la protección de derechos, definido como las restricciones y garantías requeridas para el tratamiento y protección de datos personales y privados; la transparencia y datos abiertos, que se basa en el derecho de acceso a la información pública, la divulgación constante y la publicidad de las acciones de las entidades públicas; el acceso e interoperabilidad, corresponde a las condiciones mínimas para que las entidades que desarrollen funciones públicas establezcan procedimientos que permitan el uso de datos entre sí; la eficiencia administrativa, que se refiere a la gestión de documentos generados por entidades públicas y la digitalización de información para facilitar los trámites del ciudadano; y los reportes de información, son los reportes de datos en los que las entidades privadas, la academia y la ciudadanía proveen información periódica a las entidades públicas.

Este marco jurídico se compone de disposiciones jurídicas de diferente jerarquía, que, expedidas de forma independiente para regular situaciones sociales en momentos y sectores

específicos, presenta una carencia de reglas para los riesgos asociados y una falta de articulación dentro de la generación de valor social y económico en los procesos de explotación de datos.

El PND 2014-2018 Todos por un nuevo país creó el Sistema Estadístico Nacional (SEN), el cual reúne articuladamente a las entidades, organizaciones, procesos e instrumentos técnicos para garantizar la producción y la divulgación de las estadísticas oficiales nacionales y territoriales. Este sistema es un ejemplo del aprovechamiento de datos administrativos y del intercambio de micro datos entre las entidades que hacen parte del SEN.

Estos antecedentes de política pública y de marco normativo para el aumento de la eficiencia de la gestión pública y el mejoramiento en la prestación de servicios a través del uso de las TIC, representan un aprovechamiento de datos indirecto que son la base de la generación y divulgación de información. Es decir, que las intervenciones normativas que han tenido un aprovechamiento indirecto de los datos, se puede sintetizar en: Gobierno Electrónico, eficiencia administrativa, gestión documental y protección de derechos.

En el primer tema nace en el año 2000 con el cambio de paradigma de servicios centrados en el Gobierno a la prestación de servicios al ciudadano, esto último funcionó como una estrategia que emplea las TIC para la generación de beneficios sociales. En el año 2003, el CONPES 3248 de Renovación de la Administración Pública, incluía el Gobierno Electrónico y los sistemas de información, pero se identificó una desactualización y baja calidad de la información base, la incompatibilidad de los sistemas existentes y la ausencia de un patrón que facilitara el intercambio de información. Esto conllevó a que se creara el Sistema Nacional de Información Oficial, el cual permitió crear las bases para la política de Gobierno Electrónico y para mejorar la eficiencia en la integración de las entidades públicas con los ciudadanos.

En el año 2010 en el CONPES 3650 de la Importancia estratégica de la Estrategia de Gobierno en Línea se definió el concepto de Gobierno Electrónico y con el Decreto 2693 de 2012 se modificaron algunos componentes del Gobierno en Línea como interoperabilidad, cadenas de trámites virtuales, datos y gobierno abierto. Aunque este decreto fue derogado por el Decreto 2573 de 2014, se establecieron dos instrumentos que ayudaron a la implementación de la política, estos son el Manual de Gobierno en Línea y el Marco de Referencia de Arquitectura de tecnologías de la información (TI).

En el tema de eficiencia administrativa se encuentra la normatividad relacionada con la disposición de cadenas de trámites que benefician la disminución de trámites al ciudadano y el intercambio de datos entre entidades. Las normas que han definido este tipo de intervención son: la ley 962 de 2005, que autoriza la implementación de sistemas para integrar y compartir la información; la ley 1450 de 2011, plantea el deber de las entidades públicas de compartir las bases de datos para la ejecución de sus funciones; los Decretos 235 y 2280 de 2010, establecieron que las entidades pueden usar un mecanismo viable para el intercambio de información y el Decreto 19 de 2012 denegó a las entidades solicitar información a los ciudadanos que tienen sus datos inscritos en otra entidad pública.

La gestión documental tiene disposiciones normativas que tienen la interoperabilidad como un principio en el soporte documental de cualquier entidad. La ley 594 del 2000 tiene como objetivo preservar el registro histórico de la gestión de varios documentos almacenados por su valor histórico y cultural. El Decreto 2609 de 2012, bajo reglamento de la Ley General de Archivos (594) de 2000, define que las entidades pueden implementar tecnologías y soportes nuevos para la gestión documental. El Decreto 2578 de 2012 reglamentó la creación del Sistema Integral Nacional de Archivos Electrónicos (SINAE), en el que se pone a cargo al Archivo General de la Nación el trabajo de publicar conceptos sobre el impacto de las TIC en la gestión documental para asegurar una homogeneidad en la conservación de documentos electrónicos que hacen parte del patrimonio documental digital y en la gestión de archivos del Estado.

Según el documento Base del Plan Nacional de Desarrollo 2014-2018 Todos por un nuevo país, se considera la necesidad de actualizar los archivos públicos para “avanzar en la simplificación de trámites y procesos, así como en la disminución del volumen de documentos, la reducción de costos y el incremento de la productividad de los servidores públicos.”

La materia de protección de datos personales hace referencia a la información contenida sobre una persona, este tipo de datos permite identificar a las personas e involucra derechos de habeas data, privacidad e intimidad. El Decreto 1633 de 1960 tiene en cuenta la reserva de los datos suministrados con fines estadísticos para proteger la intimidad y la información estratégica de las personas jurídicas. En 2008 se expidió la Ley 1266 para la gobernanza de la información sobre el comportamiento crediticio y financiero y los derechos asociados a los titulares.

Con la Ley 1581 de 2012 se estableció un régimen general de protección de datos personales en el que se le da la garantía a un titular de definir la recolección, uso, circulación, rectificación y eliminación de sus datos. Además, con el Decreto 1377 de 2013, se creó el Registro Único de Bases de Datos, conformado por un directorio público de bases de datos personales para identificar las canales de atención a los titulares, políticas de privacidad adoptadas, tipo de datos, medidas para garantizar su seguridad, entre otros. Asimismo, se delegó la dirección de Protección de Datos Personales de la Superintendencia de Industria y Comercio como una autoridad con amplias facultades sancionatorias.

En materia de aplicabilidad del Big Data a la Gestión de Riesgo de Desastres; De la misma forma, y bajo el acuerdo 001 de 2014 del 30 de septiembre se establecen los estatutos sobre el instituto distrital de gestión de riesgos y cambio climático: IDIGER, anteriormente conocido como FOPAE<sup>6</sup>, el cual mediante el acuerdo 546 del 2013 se transforma el Sistema Distrital de Prevención y Atención de Emergencias -SDPAE-, en el Sistema Distrital de Gestión de Riesgo y Cambio Climático-SDGR-CC, se actualizan sus instancias, se crea el Fondo Distrital para la Gestión de Riesgo y Cambio Climático “FONDIGER” y se dictan otras disposiciones.

El antiguo FOPAE, ahora IDIGER como instancia financiación y la Oficina Coordinadora para la Prevención y Atención de Emergencias según (IDIGER, 2016) se crearon en el año 1987, mediante el acuerdo 11 del mismo año; y la instancia de coordinación fue creada mediante el acuerdo 13 de 1990, y reglamentadas por medio del decreto 652 de 1990; instancias que funcionan como un solo establecimiento público, hoy en día llamado IDIGER.

Las instancias anteriormente nombradas: financiación y coordinación dependían exclusivamente del alcalde mayor hasta el año de 1995 ... “cuando la OPES fue trasladada a la Secretaría Distrital de Gobierno como UPES y posteriormente, en el año 1999 se cambió la denominación a Dirección de Prevención y Atención de Emergencias-DPAE dentro de la misma Secretaría. Mientras el FOPAE, seguía siendo un Establecimiento Público dependiendo del Alcalde Mayor.”

---

<sup>6</sup> Fondo de Prevención y Atención de Emergencias

En el Acuerdo 257 de 2006, estableció las normas básicas sobre la estructura y organización de las entidades del Distrito, se adscribe el FOPAE al sector Gobierno; y es el Decreto 413 de 2010 que modifica la estructura de la Secretaría Distrital de Gobierno, suprimió la DPAAE y todas sus funciones pasaron al FOPAE hoy en día IDIGER como Establecimiento Público.

De lo anterior, se puede concluir que la Oficina de Coordinación de Prevención y Atención de Emergencias (OPES) ha cambiado su nombre, así como su naturaleza jurídica y su ubicación dentro del organigrama del Distrito, el Fondo de Prevención y Atención de Emergencias, FOPAE hoy en día IDIGER, ha permanecido inalterado desde su creación por parte del Concejo Distrital y sólo con el Decreto 413 de 2010<sup>7</sup> se le adicionaron las funciones que hasta ese momento tenía la Dirección de Prevención y Atención de Emergencias.

De las políticas anteriormente nombradas en materia de gestión del riesgo, hacían referencia a la prevención y atención de emergencias o desastres; sin embargo más adelante dichas leyes fueron reunidas en una definición más amplia de Gestión del Riesgo de Desastre, que integra desde la prevención, la mitigación, la respuesta, preparativos y rehabilitación entre otros, en la ya nombrada ley 1523 de 2012, y fueron agrupadas en tres procesos: conocimiento del Riesgo, Reducción del Riesgo y Manejo de Emergencias, Calamidades y Desastres.

Más adelante, en el plan de desarrollo “Bogotá Humana 2012-2016”<sup>8</sup> adoptado mediante el Acuerdo Distrital 489 de 2012<sup>9</sup> y que dentro del artículo 29: programa “ Gestión Integral de Riesgos” diseña la transformación del Sistema Distrital para la Prevención y Atención de Emergencias SDPAE en: “ El sistema Distrital de Gestión del Riesgo y Cambio Climático (SDGR-CC)” unido institucionalmente y territorialmente bajo los principios de la participación, descentralización y desconcentración, con el fin de reducir y controlar los riesgos, así como los efectos del cambio climático, con el fin de manejar de manera efectiva las situaciones de desastre, calamidad o emergencia a los que Bogotá se encuentra en estado de vulnerabilidad, para este caso: *las inundaciones*:

---

<sup>6</sup>Por medio del cual se modifica el Decreto 539<sup>6</sup> de la 2006 de la Secretaría Distrital de Gobierno, se determina el objeto, la estructura organizacional y las funciones de la Secretaría Distrital de Gobierno y se dictan otras disposiciones (Alcaldía Mayor de Bogotá, 2010)

<sup>8</sup> Secretaria Distrital de Planeación, 2012-2016

<sup>9</sup> Concejo de Bogotá, ACUERDO 489 , 2012

De acuerdo con el glosario internacional de hidrología (UNESCO/OMM, 1974) la definición oficial de inundación se entiende como: “ aumento del agua por arriba del nivel normal del cauce”, así que teniendo en cuenta la anterior definición, las inundaciones son procesos que se llevan a cabo cuando existe una ocupación por parte del agua en zonas que por lo general están libres de la misma, debido comúnmente a la precipitación; es decir lluvia o granizo extremo, desbordamiento de ríos o fallas en estructuras hidráulicas o encharcamiento (CENAPRED, 2009). Sin embargo, cabe resaltar que las inundaciones y sus causas tienen gran dependencia hacia la topografía del terreno, pues también pueden ocurrir en zonas costeras debido a la alteración del oleaje, la subida del nivel del mar, monzones, entre otros; generando así daños y afectaciones en la infraestructura, cultivos y en el desarrollo de la cotidianidad de las poblaciones.

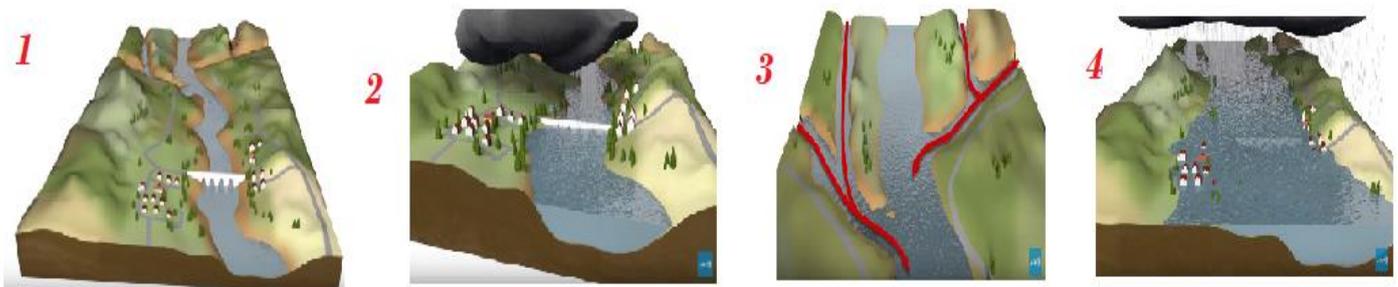
En países ubicados en zonas ecuatoriales como Colombia, las inundaciones presentan ocurrencia en los periodos más lluviosos, “generalmente en los meses donde se presenta un régimen de vientos alisios fuertes desde el oeste, donde las temperaturas ecuatoriales disminuyen y comienza la fase fría y de lluvias conocido como el Fenómeno de la Niña” (IDEAM, 2012), no obstante, las inundaciones también pueden ser causadas por sucesos climáticos anormales, ajenos a eventos de mayor duración e intensidad como el fenómeno de la Niña anteriormente nombrado.

Así mismo, existen causas antrópicas que pueden llegar a aumentar el riesgo de inundaciones, puesto que actividades humanas como: ...la urbanización, provoca que el suelo se cubra con materiales poco permeables que no permiten que el agua se filtre en el suelo de manera regular, así mismo la deficiencia en el alcantarillado y la falta de cultura ciudadana que taponan las alcantarillas de basura, evitan que dichas sean capaces de conducir grandes volúmenes de agua; (CENAPRED, 2009) la tala de árboles también es una causa antrópica de estrecha relación con las inundaciones pues al llevar a cabo esta actividad, se destruye la capa vegetal del suelo, dejando así el terreno expuesto a ser arrastrado en eventos de lluvias, desde las zonas más altas hacia las zonas más bajas, taponando los drenajes y obstruyendo ríos y canales lo que favorece e incrementa la ocurrencia de inundaciones.

Los tipos de inundaciones más comunes son: los pluviales y los fluviales; las primeras ocurren cuando el terreno se ha saturado y se lleva a cabo un proceso de saturación del agua excedente, como resultado de largos periodos de lluvia, así como los diferentes factores anteriormente

nombrados, que pueden alterar la permeabilidad del terreno; las segundas se generan cuando el agua se desborda de los ríos y se ubica en los terrenos aledaños a dichos causes.

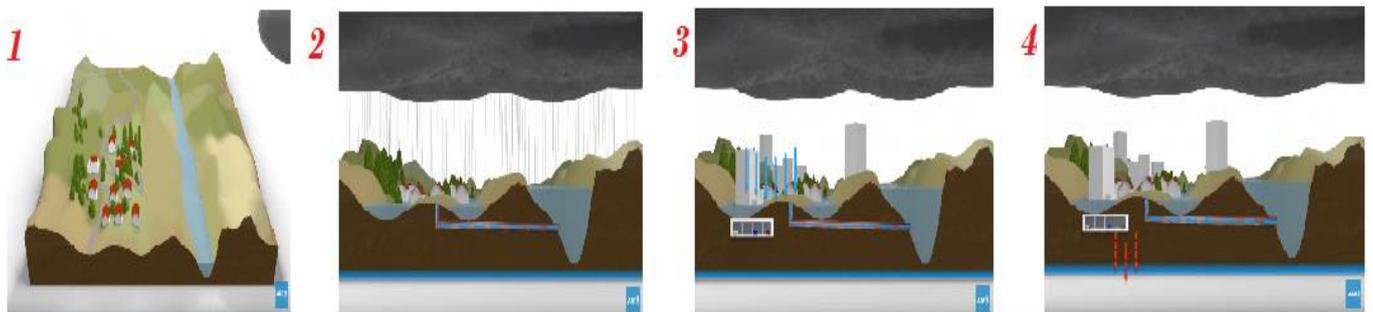
*Figura 2. Inundaciones Fluviales*



Fuente de las imágenes: (D'epenoux, KOGUC, & Hennebert, 2018): <https://www.youtube.com/watch?v=Nov4WSezIMU>

La figura 2 se observa que este tipo de inundación se produce por: 1 un desborde, 2. luego de fuertes precipitaciones un río puede salir de su cauce inundando zonas enteras alejadas de las zonas de lluvias; 3. así mismo los ríos tributarios pueden empeorar las cosas al aumentar el caudal de agua empeorando las 4 inundaciones río abajo.

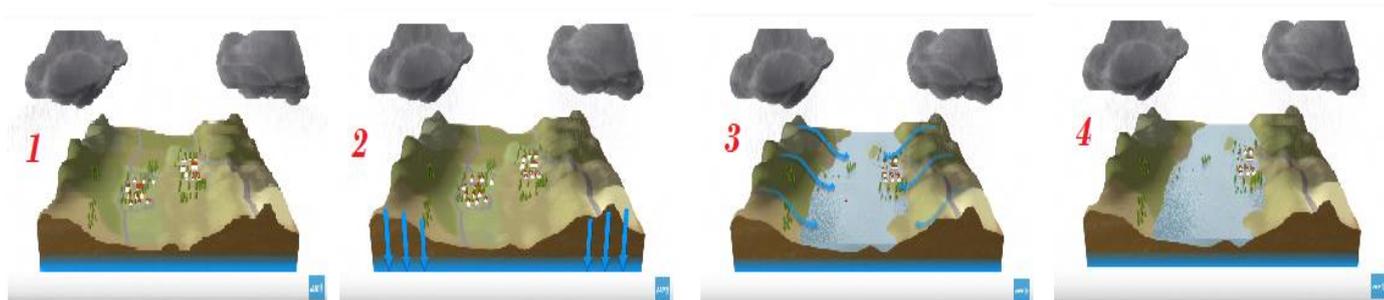
*Figura 3. Inundaciones Urbanas*



Fuente de las imágenes: (D'epenoux, KOGUC, & Hennebert, 2018): <https://www.youtube.com/watch?v=Nov4WSezIMU>

La figura 3 muestra que las inundaciones urbanas: 1 se deben a la creciente 2 urbanización de zonas inundables, en donde el agua sube mediante los desagües en las zonas más bajas, de igual manera la 3 construcción de edificios y superficies impermeables como el concreto empeoran el fenómeno, ya que menos agua puede ser absorbida por el suelo, 4 sobrecargando así al sistema de alcantarillado; estas inundaciones son por encharcamiento.

Figura 4. Inundaciones Pluviales



Fuente de las imágenes: (D'epenoux, KOGUC, & Hennebert, 2018): <https://www.youtube.com/watch?v=Nov4WSezIMU>

La figura 4 se evidencia que: 1 es el resultado de acumulación de agua en las cuencas, en donde bajo condiciones climáticas normales 2 la mayor parte del agua es absorbida por el suelo en donde confluye con la capa freática, 3 sin embargo cuando el suelo está saturado, 4 el agua se acumula en las zonas más bajas.

En Colombia, las emergencias por inundaciones están dadas por la combinación de diferentes factores que además de fluctuaciones normales en el estado del tiempo, alteran e interfieren con las dinámicas naturales del territorio: la falta de ordenamiento territorial, la deforestación y la descontrolada expansión de la frontera agrícola son algunas de las causas que contribuyen a que fenómenos como las inundaciones, dejen como resultado grandes afectaciones cada año, donde en temporada invernal miles de familias se ven afectadas por el desbordamiento de quebradas, ríos, lagunas o por el taponamiento de alcantarillado que impide el normal flujo de desagüe de las aguas lluvia, que terminan reflejando pérdidas económicas (viviendas, enseres, cultivos y animales).

Las inundaciones más fuertes que se han presentado en Colombia por fenómeno de la niña fue durante el periodo 2010-2011, dicho fenómeno comenzó su etapa de formación hacia el mes de junio de 2010 y encontró su máximo fortalecimiento alcanzando su etapa de madures en enero del 2011, en donde según (IGAC, IDEAM, & DANE, 2011) se categorizó como fuerte.

Mediante el análisis ONI<sup>10</sup> ; el IDEAM pudo establecer un escenario de inundaciones así como de las zonas del país más vulnerables y afectadas por inundaciones:

Tabla 1. Análisis ONI para escenarios de inundaciones.

AÑO	EFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDE	DEF
1990	0,1	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,3	0,3	0,3	0,4
1991	0,4	0,3	0,3	0,4	0,6	0,8	1,0	0,9	0,9	1,0	1,4	1,6
1992	1,8	1,6	1,5	1,4	1,2	0,8	0,5	0,2	0,0	-0,1	0,0	0,2
1993	0,3	0,4	0,6	0,7	0,8	0,7	0,4	0,4	0,4	0,4	0,3	0,2
1994	0,2	0,2	0,3	0,4	0,5	0,5	0,5	0,6	0,6	0,7	0,9	1,3
1995	1,2	0,9	0,7	0,4	0,3	0,2	0,0	-0,2	-0,5	-0,6	-0,7	-0,7
1996	-0,7	-0,7	-0,5	-0,3	-0,1	-0,1	0,0	-0,1	-0,1	-0,2	-0,3	-0,4
1997	-0,4	-0,3	0,0	0,4	0,8	1,3	1,7	2,0	2,2	2,4	2,5	2,5
1998	2,3	1,9	1,5	1,0	0,5	0,0	-0,5	-0,8	-1,0	-1,1	-1,3	-1,4
1999	-1,4	-1,2	-0,9	-0,8	-0,8	-0,8	-0,9	-0,9	-1,0	-1,1	-1,3	-1,4
2000	-1,6	-1,4	-1,0	-0,8	-0,6	-0,5	-0,4	0,4	-0,4	-0,5	-0,6	-0,7
2001	-0,6	-0,5	-0,4	-0,2	-0,1	0,1	0,2	0,2	0,1	0,0	-0,1	-0,1
2002	-0,1	0,1	0,2	0,4	0,7	0,8	0,9	1,0	1,1	1,3	1,5	1,4
2003	1,2	0,9	0,5	0,1	-0,1	0,1	0,4	0,5	0,6	0,5	0,6	0,4
2004	0,4	0,3	0,2	0,2	0,3	0,5	0,7	0,8	0,9	0,8	0,8	0,8
2005	0,7	0,5	0,4	0,4	0,4	0,4	0,4	0,3	0,2	-0,1	-0,4	-0,7
2006	-0,7	-0,6	-0,4	-0,1	0,1	0,2	0,3	0,5	0,6	0,9	1,1	1,1
2007	0,8	0,4	0,1	-0,1	-0,1	-0,1	-0,1	-0,4	-0,7	-1,0	-1,1	-1,3
2008	-1,4	-1,4	-1,1	-0,8	-0,6	-0,4	-0,1	0,0	0,0	0,0	-0,3	-0,6
2009	-0,8	-0,7	-0,5	-0,1	0,2	0,6	0,7	0,8	0,9	1,5	1,5	1,8
2010	1,7	1,5	1,2	0,8	0,3	-0,2	-0,6	-1,0	-1,3	-1,4	-1,4	-1,4
2011	-1,3	-1,2	-0,9	-0,6	-0,2	0,0						

Serie ONI (°C) 1990 a 2011

✓ Con anomalías positivas (mayores a 0.5°C) que persistan más de cinco periodos se identifican eventos El Niño.

✓ Con anomalías negativas (inferiores a -0.5°C) que persistan más de cinco periodos se identifican eventos La Niña

Tomado de: National Oceanographic Atmospheric Administration, Climate Prediction Center; (IGAC, IDEAM, & DANE, 2011)

Como se observa en la anterior serie ONI para el periodo 2010-2011 se presentó fenómeno de la niña en donde se produjeron fuertes aguaceros que causaron emergencias por inundaciones en diferentes zonas del país; en donde en para el periodo de gobierno correspondiente en cabeza del ex presidente Juan Manuel Santos, se declaró emergencia Económica, Social y Ecológica<sup>11</sup>, así como situación de desastre con el fin de hacerle frente a las inundaciones y deslizamientos producidos por la temporada invernal.

Para el año 2011 según el DANE, dentro de los resultados de registro único de Damnificados por la emergencia invernal 2010-2011, que se realizó con el fin de identificar y caracterizar a la población Damnificada por la emergencia invernal, con el objetivo específico de orientar los

<sup>10</sup>Artículo científico de la revista ClimateWatch Magazine. Recuperado de: <https://www.climate.gov/news-features/understanding-climate/climate-variability-oceanic-ni%C3%B1o-index>

<sup>11</sup> Decreto 020 del 7 de enero de 2011 (Sentencia C-216 de 2011 Corte Constitucional, 2011)

proyectos y programas del Gobierno Nacional para dicha población, en cuento a ayudas humanitarias, reconstrucción y rehabilitación de las zonas afectadas, las cifras fueron:

*Tabla 2. Cifras de la emergencia invernal 2010-2011*

Cifras de la emergencia invernal 2010-2011	
Total de municipios afectados <sup>12</sup>	998
Hogares potencialmente Damnificados <sup>13</sup>	647.017
Hogares potencialmente afectados	232.525
Personas potencialmente damnificadas	2.350.207
Personas potencialmente afectadas	869.032

Fuente de los datos: (IGAC, IDEAM, & DANE, 2011)

Para atender la emergencia por la ya nombrada ola invernal que afecto a Colombia entre los periodos 2010-2011, se ideó un plan de contingencia por parte del ministerio de ambiente, vivienda y desarrollo territorial, en cabeza de Beatriz Uribe Botero, se destinó ayudas a personas que se encontraban en alto riesgo no mitigable, para las cuales se solicitaron recursos adicionales para hacerle frente a situaciones adversas (Minambiente, 2010), en donde después de totalizar la cantidad de pérdidas materiales y humanas por parte del DANE<sup>14</sup>, se asignó un presupuesto de aproximadamente \$1.80 billones de pesos Colombianos, en donde según él (Minambiente, 2010) las fuentes de financiación se distribuyeron de la siguiente forma: “\$631.261 millones de pesos provinieron del subsidio familiar de vivienda, así como del fondo nacional de vivienda, \$564.369 millones de pesos provinieron de las diferentes entidades territoriales es decir, municipios y departamentos, y finalmente \$ 564.369 millones de pesos, que fueron adicionales del fondo nacional de regalías y presupuesto nacional.”

Durante y después de la ola invernal se desataron diferentes culpabilidades, y se responsabilizó a las corporaciones autónomas regionales pues se consideró que no se habían desarrollado suficientes actividades de prevención en orden de minimizar los impactos y efectos de las

<sup>12</sup> Aquel que sufre daños menores por efecto de un desastre de origen natural o inducido por actividad humana y que podría o no requerir ayuda humanitaria. (INDECI, 2019)

<sup>13</sup> Aquel que sufre daños severos por un desastre de origen natural o inducido por actividad humana y que requiere ayuda humanitaria inmediata. (INDECI, 2019)

<sup>14</sup> Sistema de Consulta Registro Único de Damnificados -REUNIDOS- Tomado de: <http://sige.dane.gov.co/reunidos/>

inundaciones; a razón de lo anterior el Gobierno Nacional decidió tomar diferentes acciones entre las que se destacan la creación de nuevas corporaciones, así como la fusión de otras, no obstante estas acciones se declararon inexequibles por la corte constitucional mediante el decreto 141 de 2011 (Decreto 141 de 2011 Nivel Nacional).

A razón de lo anterior, el gobierno creó la ley 1523 del 2012 ( Congreso de la Republica, 2012) por la cual se adopta la política nacional de gestión del riesgo de desastres y se establece el Sistema Nacional de Gestión del Riesgo de Desastres<sup>15</sup>, en donde define a la gestión del riesgo como un proceso social que se orienta hacia la formulación, ejecución, seguimiento y evaluación de diferentes políticas, estrategias, planes, programas, proyectos, regulaciones e instrumentos, así como acciones permanentes necesarias para el conocimiento y la reducción del riesgo y el manejo de desastres, con el objetivo de contribuir al bienestar, la seguridad, el desarrollo sostenible y la calidad de vida de las personas. Así mismo, esta política se enmarca como una política indispensable para garantizar los derechos colectivos, la seguridad territorial y social, la sostenibilidad de los territorios y mejorar la calidad de vida de las comunidades y poblaciones expuestas a cualquier tipo de riesgo; por lo anterior está estrechamente relacionada con la planificación y desarrollo seguro del territorio, que con junto la gestión territorial sostenible, se enfoca en casi todos los niveles de gobierno para la efectiva participación y protección de la población.

El IDEAM, como institución pública de apoyo técnico y científico al Sistema Nacional Ambiental: definió a nivel nacional las áreas inundables en relación a las áreas de los cuerpos de agua, y se identificó que uno de los factores que afecta y determinan los escenarios de inundación es generado por la variabilidad y el cambio climático; de esta manera el IDEAM ha ido monitoreando la susceptibilidad a inundaciones<sup>16</sup>, eventos de inundación<sup>17</sup>, entre otros.

De igual manera el IDEAM a través de la plataforma Fews Colombia ( sistemas de pronósticos hidrológicos y alertas tempranas), se ha venido trabajando en la modernización y actualización

---

<sup>15</sup> El objetivo general de la entidad es : Llevar a cabo el proceso social de la gestión del riesgo con el propósito de ofrecer protección a la población en el territorio colombiano, mejorar la seguridad, el bienestar y la calidad de vida y contribuir al desarrollo sostenible. ( Congreso de la Republica, 2012)

<sup>16</sup> Indica el tipo de inundación (lenta o súbita) y los límites externos para un evento extremo. Es el primer paso para elaborar mapas más detallados.

<sup>17</sup> Muestra los eventos de inundaciones observadas y el registro de su extensión de inundación (límites, nivel registrado, caudales, profundidades y secciones transversales).

de alertas tempranas por inundación, mediante la implementación del anteriormente nombrado sistema operacional de pronóstico, “el cual está basado en el software Delf-FEWS desarrollado por Deltares (instituto de investigación holandés) y donada al IDEAM como parte de una Cooperación Internacional entre Los Países bajos y Colombia.” (IDEAM, 2019), que le permite al IDEAM realizar una adecuada gestión de los procesos de pronóstico hidrológico, como parte de una modelación hidrológica, hidráulica y estadística, así mismo integra datos de series de tiempo de diferentes y formatos tanto del IDEAM como de corporaciones autónomas regionales;... [Desde el 2014 a la fecha se han acoplado alrededor de 20 modelos, entre hidráulicos, hidrológicos y estadísticos. Estos modelos proveen el pronóstico hidrológico en más de 90 de puntos de monitoreo del IDEAM, CAR y CVC con un horizonte de pronóstico de 3 días. Estos modelos generan pronóstico para los cauces principales de los ríos Magdalena, Cauca, Meta, San Jorge, Lebrija, La Vieja, ente otros.]

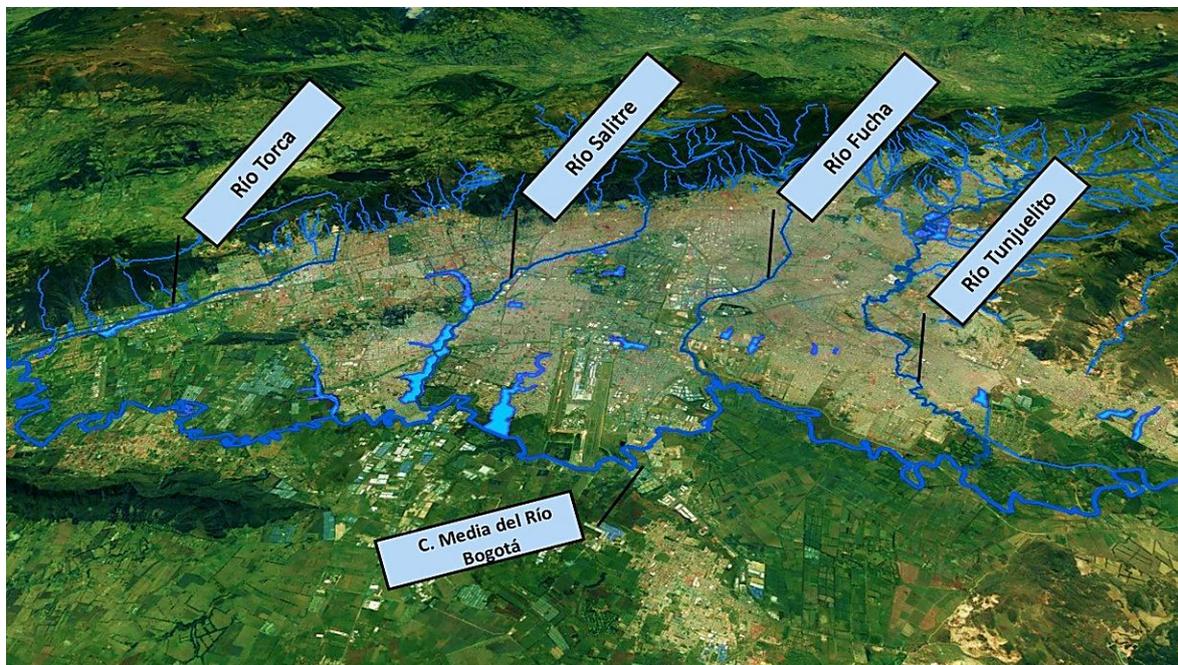
Bogotá como distrito capital es la ciudad más poblada de Colombia con casi ocho millones de personas (DANE, 2018), por lo tanto, también es una de las ciudades que cuenta con el mayor porcentaje de área construida que según él (Banco Mundial ,2012) corresponde al 70% de las capitales del país; tiene aproximadamente 300 kilómetros cuadrados de área urbana y una densidad poblacional de 4.146 habitantes por kilómetro cuadrado; si se compara las cifras del censo del año 2005 con las nuevas cifras del boletín preliminar del censo 2018, Bogotá ha ganado un poco más de un millón quinientos mil habitantes en la última década, debido a los múltiples conflictos sociales del país que han generado altos índices de migración interna hacia la capital debido esencialmente al gran índice de desarrollo y oportunidades económicas, ya que según (DIAN, 2015): Bogotá presenta las mejores oportunidades de trabajo y servicios.

Lo anterior, incrementa una dinámica de uso de suelo que impulsa a la creación y construcción de grandes superficies impermeables en el suelo que son típicas de las grandes ciudades, y ciudades en proceso de urbanización; que son el claro resultado de la evolución de los procesos de asentamientos urbanos desordenados que generan deterioro en la calidad y prestación de servicios básicos públicos y de saneamiento, así como la destrucción, compactación, relleno y pavimentación de los humedales y crecimiento desmesurado y exponencial de asentamientos en zonas susceptibles a inundaciones; en donde según el IDIGER existen más de 450 zonas en donde no se cumplen las rondas mínimas de los cuerpos de agua superficiales enmarcados en el artículo 78 del decreto distrital 190 de 2004, ya que dichas rondas están siendo invadidas con construcciones y rellenos producto del depósito de escombros, lo que altera la dinámica de los

cauces, lo que a su vez disminuye la infiltración de agua lluvia por parte del suelo, aumentando la escorrentía y disminuye los tiempos de concentración, que deja como resultado un incremento en la amenaza de inundaciones por encharcamiento en la ciudad.

La ciudad de Bogotá, se encuentra dividida en cuatro cuencas o drenajes correspondientes a los ríos: Salitre, Fucha, Tunjuelo que fluyen de oriente a occidente originándose mayormente en los cerros orientales; y el río Torca que fluye de sur a norte, son los principales tributarios de la cuenca media del río Bogotá; como se observa a continuación:

*Figura 5. Zonificación hidrográfica de Bogotá*



Fuente: IDECA e IDIGER

Las principales causas de inundación se deben a lluvias fuertes y de gran duración que caen sobre la ciudad y generan volúmenes de agua que los ríos, quebradas y humedales no tienen la capacidad para recolectar y transportar, sumado a mala disposición en los cauces de residuos sólidos y escombros y rompimiento de los jarillones que se construyen a lo largo de los ríos.

Bogotá se encuentra ubicada en la región Andina, la cual posee una gran diversidad pluviométrica, con lluvias escasas con valores que van hasta 2000 mm a lo largo de la cordillera oriental (Arango, C., Dorado, D., & Ruiz, 2012); Bogotá presenta valores que oscilan entre los 700 mm a 1000 mm totales anuales en donde los periodos con menor ocurrencia son: enero, febrero y entre julio y agosto con valores de 32 mm a 65 mm y 33 mm a 48 mm,

respectivamente; mientras que los meses de abril a mayo y octubre a noviembre los valores de precipitación varían desde 92mm a 142mm.(IDEAM, 2015).

La sabana de Bogotá D.C. se caracteriza por presentar un régimen climático bimodal donde según (IDIGER, 2018) se caracteriza por tener “entre 150 a 200 días con lluvia en todo el año, se aprecia un período con menor cantidad de días lluviosos, ubicado en enero y febrero y un segundo período, mucho menos pronunciado, en julio y agosto. En el primer período llueve entre 9 y 10 días/mes; en el segundo, entre 15 y 16 días. Los demás meses presentan más eventos de precipitación tienen cerca de 17 días lluviosos con un máximo de 19 días, especialmente en el segundo periodo en el mes de octubre”.

De lo anterior, se puede inferir que Bogotá es una ciudad que presenta altos índices de precipitación, que junto con los altos índices de deforestación de las cuencas principales que la atraviesan (ver figura 5), sumado a los procesos de impermeabilización como consecuencia de la continua expansión urbana que ha venido presentando Bogotá en los últimos años (Beltrán, 2019), afecta el ciclo hidrológico y disminuye la capacidad de absorción del agua en el suelo, incrementando los niveles en la precipitación neta<sup>18</sup> provocando así que los flujos superficiales por escorrentía entren de forma directa al caudal de los ríos urbanos ( Fucha, Salitre, Tunjuelo y Torca), así como a los sistemas de drenajes y aguas lluvias; según (IDIGER, 2018): En ciudades como Bogotá con alto desarrollo urbano e incidencia de la variabilidad climática ocasionan que la precipitación sea la principal variable de estudio y que junto con otros factores físicos de las cuencas sean el principal detonante de fenómenos de inundaciones por encharcamiento.

En Bogotá la amenaza por inundación se debe principalmente al desbordamiento y encharcamiento de los causes que discurren sobre la planicie inundable del río Bogotá y en la Sabana de Bogotá; por lo tanto incluye los causes: principal del río Bogotá, y los afluentes anteriormente nombrados en su zona baja.<sup>19</sup> Las pendientes más bajas es decir las que son menores a 0.5% se encuentran en las zonas de expansión urbana; también las diferentes

---

18 Es la lluvia que causa escorrentía superficial. Es igual a la lluvia total menos la intercepción menos la detención superficial y menos la infiltración (Universidad de Salamanca España, 2019).

19 La diferenciación de estos sectores es básicamente debido a la pendiente de los ríos, que debe estar por debajo de 0.5%, para presentar flujo subcrítico y por lo tanto condiciones de remanso tipo M1 o M2 durante las avenidas mayores, condiciones bajo las cuales existen los desbordamientos normales sobre la planicie. (IDIGER, 2018)

intervenciones hechas por parte de la empresa de Acueducto y Alcantarillado de Bogotá afectan las condiciones hidráulicas del mismo.

La primera intervención drástica que se le realizó al río Bogotá fue la construcción de las compuertas de Alicachín<sup>20</sup> que se ubicó al finalizar la sabana de Bogotá, aguas arriba del Salto de Tequendama con el fin de mantener los niveles de caudal en esta zona para darle paso a la desviación de los caudales hacia la llamada entonces planta de Canoas<sup>21</sup> Según el IDIGER: “Esta intervención, ha limitado fuertemente el propio desarrollo morfológico del río Bogotá, y desde luego afecta los niveles de agua en un trecho bastante grande aguas arriba, en forma tal que los niveles resultan siempre superiores a los que naturalmente generaría la corriente en su condición natural”. El anterior efecto sobre el cauce del río se ha extendido aproximadamente hasta la desembocadura del río Tunjuelo, y se ha convertido en otra de las agravantes principales de las inundaciones en Bogotá.

Con el fin de establecer las condiciones de inundación en la zona urbana de Bogotá, el IDIGER llevo a cabo una serie de estudios de zonificación, amenaza, geotecnia, suelos, entre otros, y con base en dichos estudios elaboro el mapa de zonificación de amenaza por inundación de Bogotá, identificando las zonas con amenaza: alta, media y baja.

La figura 6 muestra que la amenaza alta corresponde a la zona delimitada por la línea de inundación producida por el desborde del cauce, para un caudal con un periodo de retorno menor o igual a 10 años, sin discernir las causas (naturales o antrópicas intencional y no intencional), con una profundidad en la lámina de agua igual o superior a 0.50 m, y con una probabilidad de ocurrencia mayor al 65%. Esta franja tiene una probabilidad de estar inundada por lo menos una vez cada diez años durante la vida útil del jarillón hasta ese nivel. La amenaza media está calculada para un caudal con periodo de retorno de 10 a 100 años, con una profundidad en la lámina de agua de más de un metro de altura y con probabilidad de ocurrencia entre 10% a 65%; finalmente la amenaza baja es la zona delimitada por la línea de inundación generada por el desborde del cauce calculado para el caudal de agua creciente, en un periodo de retorno mayor o igual a 100 años, con una lámina de agua superior a un metro con efectos potencialmente dañinos leves para la comunidad; dicha franja tiene la probabilidad de

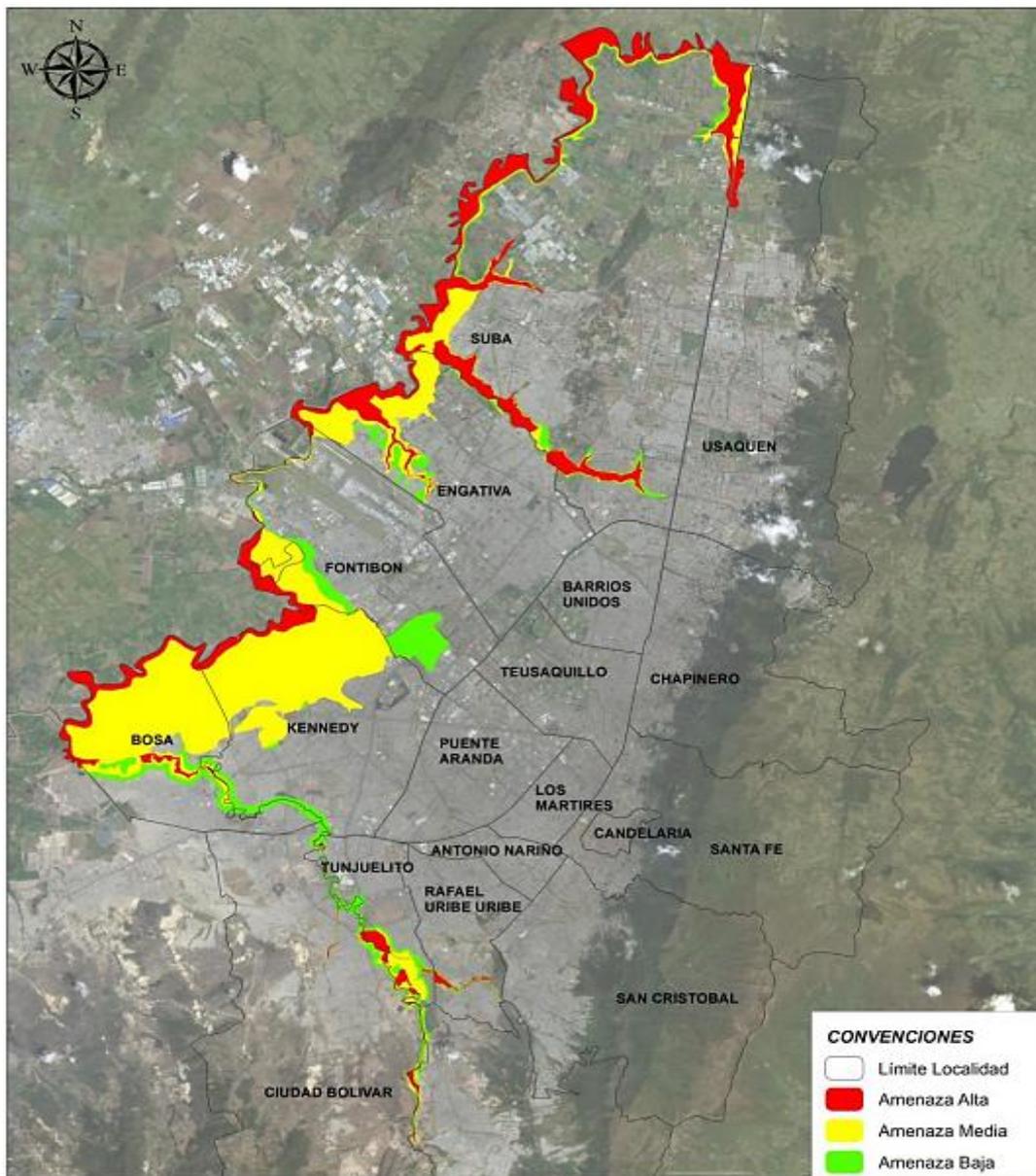
---

20 Corporación Autónoma Regional (CAR, 2014)

21 Proyecto estructuración PTAR Canoas. Recuperado de: <https://www.wsp.com/es-CL/proyectos/ag-estructuracion-ptar-canoas>

inundarse por lo menos una vez cada 100 años con una probabilidad de ocurrencia menor al 10%.

Figura 6. Zonificación de amenaza por inundación en Bogotá



Fuente: IDIGER

Debido a lo anterior, Bogotá se ha destacado en los últimos años por su actitud proactiva frente a la gestión integral del riesgo, específicamente frente a la reducción de la vulnerabilidad y en conocimiento de la amenaza, como componentes principales del desarrollo sostenible, sin embargo, se ha dejado un poco de lado las acciones frente a la atención a desastres.

Así mismo, la creación de diferentes estudios frente a amenazas a un evento adverso que afecte potencialmente la ciudad de Bogotá, como lo son las inundaciones, están orientados a reducir la vulnerabilidad de la ciudad frente a eventos de emergencia ya que estiman la intensidad de los mismos y los compara con la variable de estudio que refleje la severidad de dicha emergencia con el objetivo de poder cuantificar, medir y comparar, y poder predecir lo que posiblemente pueda ocurrir dentro de los diferentes rangos de confiabilidad de la lámina de agua, su velocidad en las zonas de tránsito de los cauces, y predecir de forma efectiva la altura de la lámina de agua del evento adverso.

Gracias a los anteriores estudios se han podido ir desarrollando algunos sistemas distritales o municipales de prevención de desastres y alertas tempranas, que integran metodologías, formatos y generan alimento informático de bases de datos que ayudan a describir el tipo de evento, duración e intensidad así como su distribución espacial que ayudan y tienen como objetivo describir la afectación sobre el ambiente, la infraestructura y la vida tanto humana como animal, y permiten generar un perfil adecuado que apoye la definición de políticas de prevención, planes y acciones para la prevención y reducción de afectaciones por fenómenos naturales como las inundaciones.

### 3.2 Marco conceptual

#### 3.2.1 Big Data

Son grandes conjuntos de datos que superan la capacidad de ser capturados, procesados y gestionados por las tecnologías de la información tradicionales. Estos se producen a través de sensores, dispositivos y la actividad de los usuarios de Internet a una alta velocidad, para que la información obtenida permita tomar decisiones correctas en cualquier momento. Las características fundamentales del Big Data son:

- Volumen: hace referencia a la cantidad de datos. Es decir, que los datos relacionados al Big Data se generan en mayores cantidades que los datos tradicionales. Las unidades de medida para estos datos son del orden de terabytes (un billón de bytes), petabytes (mil billones de bytes), exabytes (un millón de billones de bytes), zettabytes y otros. Según Joyanes (2013) en Twitter se producen más de 9 terabytes (TB) de datos por día,

Facebook 10 TB y algunas empresas forman terabytes de datos cada hora de todos los días del año.

- Veracidad: es la fiabilidad de los datos. Este aspecto supone un gran reto, debido a que se manejan diversos tipos y fuentes de datos en la toma de decisiones.
- Velocidad: se le atribuye al aumento creciente de los flujos de datos, o sea la rapidez en la que se generan y se procesan los datos en su entorno de explotación. Hay que tener en cuenta que los datos se producen de manera continua y en tiempo real para mejorar la toma de decisiones sobre una base de información original. Por ejemplo Twitter permite descargar tweets en tiempo real y reconocer los “trending topics” de los temas que son tendencia en cualquier momento y lugar.
- Variedad: tiene que ver con la gestión de diversas fuentes, tipos y formatos de datos, como los datos estructurados, semiestructurados y no estructurados.

Los datos estructurados tienen un formato con campos fijos. En estas fuentes, los datos son detallados y conforman las bases de datos relacionales, hojas de cálculo y archivos principalmente. Además, estos datos se constituyen de piezas de información que vienen en un formato específico y en un orden concreto; algunos ejemplos de estos formatos pueden ser: fechas, números de documentos, números de cuentas en un banco y otros.

Los datos semiestructurados poseen un flujo lógico y un formato que tiende a definirse, pero su comprensión resulta ser un poco difícil por el usuario. A pesar de que estos datos no contienen formatos fijos, existen etiquetas y marcadores que separan los elementos del dato. Un ejemplo de datos semiestructurados son los registros web logs de las conexiones a Internet, los cuales tienen diferentes piezas de información que sirven para determinado propósito, un ejemplo de esto, es el texto etiquetas de lenguajes HTML u XML.

Los datos no estructurados son datos que se almacenan como objetos sin una estructura uniforme y con poco o ningún control sobre ellos. Los datos de texto, fotografía, audio, video, cartas, imágenes digitales, mensajes de texto, mensajes de correo electrónico,

artículos, libros, mensajes de mensajería instantánea y demás. En esta fuente, el 80% de la información de las organizaciones no se almacena en bases de datos relacionales sino que se encuentra completamente esparcida.

- Valor: Es una forma eficiente y económica de aprovechar los datos mediante algoritmos analíticos de gran rendimiento y herramientas para la visualización. Esto quiere decir que el valor económico se da cuando los datos se convierten en información con el fin de generar un conocimiento útil para la acción o toma de decisiones. (Instituto Valenciano de Tecnologías Turísticas, 2015; Gutiérrez Puebla, 2018; Aguilar, 2013; Marqués, 2015).

### 3.2.2 Redes sociales

Las redes sociales son sitios en la red que facilita a los usuarios la interacción, comunicación, distribución de contenido y creación de comunidades. Algunas características que se destacan de las redes sociales es que permiten al usuario crear un perfil público o semi-público en un sistema cerrado, articulan listas de distintos usuarios que comparten una conexión y permiten visualizar y explorar su propia lista de contactos junto con las que son creadas por otros usuarios. (Cardoso et al., 2012; Cueto, José, Corzo, & Joseph, n.d.; Rica et al., 2012).

### 3.2.3 Inundaciones

Son fenómenos hidrológicos producidos por una acumulación excesiva de agua fuera de su cauce natural o artificial. Se caracterizan por ser potencialmente destructivos y originados por lluvias constantes y generalizadas, las cuales van aumentando el nivel de las aguas de los cauces de forma progresiva, hasta superar la capacidad de retención o de transporte de los canales y ocasionar un desbordamiento de las aguas sobre las llanuras de inundación.

Estos tipos de eventos son periódicos y típicos de la dinámica natural de las cuencas hidrográficas. Se pueden clasificar en dos tipos: inundaciones lentas o de tipo aluvial y crecientes súbitas o inundaciones de tipo torrencial. Las primeras ocurren en las zonas llanas de los ríos y con valles aluviales amplios, el nivel aumenta diariamente unos cuantos centímetros, produce afectaciones de considerables extensiones, pero hay pocas pérdidas de vidas. La afectación puede llegar a durar algunos meses.

Las crecientes súbitas tienen un mayor poder destructivo, a pesar de que las áreas de afectación son menores, también hay mayor pérdida de vidas, se generan de forma rápida dependiendo la ocurrencia de precipitaciones en las partes altas de las cuencas, los niveles de agua aumentan en metros dentro de pocas horas. La permanencia de estas inundaciones en las zonas afectadas es de horas o pocos días y generalmente se presentan en las cuencas de pendiente alta, como las de la región Andina. (UNGRD, 2018; IDEAM, s.f.; Velandia, 2014).

#### 3.2.4 Desastre

Es una alteración grave generada por el impacto de un fenómeno natural o antropogénico que interrumpe el funcionamiento cotidiano de una comunidad. Tiene como causa una combinación de amenazas, condiciones de vulnerabilidad y medidas insuficientes para disminuir el riesgo y como consecuencias daños, pérdidas humanas, económicas, ambientales y materiales. Así mismo se caracteriza por ser un evento inesperado, por involucrar experiencias inusuales y producir una sensación colectiva de inseguridad, peligro real o imaginario.

El desastre hace que se exijan acciones por parte del Estado y del sistema nacional para la respuesta adecuada ante una emergencia, la rehabilitación y reconstrucción. Es importante que cualquier sociedad tenga políticas de gestión de riesgos para evitar los efectos de los desastres mediante la identificación de las condiciones de amenaza y vulnerabilidad, lo que permitirá contar con información orientada a planificar y tomar decisiones para reducir el riesgo. (Ballis, s.f.; EIRD y OCAH, 2008; Ley 1523, 2012; SAV y DGPC, 2015).

#### 3.2.5 Respuesta

Ejecución de acciones, provisión de ayuda o intervención dada durante o después de un evento adverso para preservar vidas, cubrir las necesidades básicas de la población afectada, disminuir pérdidas y reducir el sufrimiento. Estas acciones se presentan en aspectos como: el transporte, telecomunicaciones, evaluación de daños, análisis de necesidades, salud, saneamiento básico, búsqueda y rescate, extinción de incendios, servicios públicos, albergues y alimentación, seguridad, convivencia, financiamiento, manejo general de la respuesta, entre otros. La respuesta puede ser inmediata, a corto plazo o prolongada y su efectividad depende de la calidad de preparación. (SAV y DGPC, 2015; Ley 1523,2012; EIRD y OCAH, 2008).

### 3.3 Estado del arte o del conocimiento

En primer lugar se tiene que el caso del estado de Tabasco en México, presentado dentro del marco del proyecto: *utilización de la actividad de telefonía móvil para la gestión de desastres durante las inundaciones (using mobile phone activity for disasters management during floods)*, del Pulso Global de las Naciones Unidas: Aprovechar los grandes datos para el desarrollo y la acción humanitaria (UNITED NATIONS GLOBAL PULSE: harnessing big data for development and humanitarian action ), presentado por Pastor- Escuredo, D., Morales-Guzmán, A. et, al bajo el nombre de : *Flooding through the lens of mobile phone activity (Inundaciones a través de la lente de la actividad del teléfono móvil)*, y publicado posteriormente por el Banco Interamericano de desarrollo en el Blog sobre conocimiento abierto en América Latina y el Caribe.

Según el Programa de las naciones unidas para el desarrollo (PNUD, 2014) “El estado de Tabasco en México entre los años 2007 y 2011, sufrió numerosas inundaciones que generaron daños y pérdidas por más de 57 millones de pesos (3 Millones USD).

Para el caso de estudio en general se tomó el año 2009 en donde UN Global Pulse junto con investigadores de la Universidad Politécnica de Madrid (UPM) y Telefónica Research realizaron un proyecto para explorar el uso de Big Data como herramienta para acción humanitaria; junto con el asesoramiento por parte de expertos del Gobierno de México y el Programa Mundial de Alimentos (PMA); el objetivo del proyecto era determinar si mediante la utilización de Big Data era posible crear nuevos mecanismos de alerta temprana y gestión de inundaciones.

En el estudio se comparó la reacción social, entendidos como los patrones del Big Data; la actuación y reacción de las instituciones y finalmente el patrón de lluvias, como factor externo; “Para medir la reacción social se analizaron registros anonimizados de llamadas de móviles, asegurando la privacidad de usuarios. La actuación institucional se recuperó mediante noticias, registros públicos y actas de protección civil sobre las inundaciones de 2009 y 2010 en México. Para medir el avance de las inundaciones se utilizaron datos abiertos de satélites de la NASA.” (Pastor- Escuredo, D., Morales-Guzmán, A. et, al, 2014).

El caso de Tabasco destaca como ejemplo en la utilización de datos masivos Big Data, para mejorar la respuesta a desastres naturales, y demostró que aunque se debe tener precaución al momento de anticipar un desastre natural real , para este caso una inundación, si es posible localizar las zonas más afectadas de forma rápida y eficiente, así como medir los intervalos de tiempo entre la ocurrencia del fenómeno propio y periódico de la dinámica natural de las cuencas hidrográficas sumados con fenómenos climáticos, los efectos y las acciones encaminadas a la mejora de la gestión.

En esta misma línea de investigación, se encuentra el *Big Data como aliado para afrontar catástrofes naturales como el ciclón australiano Debbie*, publicado por Banco Bilbao Vizcaya Argentaria, S.A (BBVA); Según (Romero, 2017) actividades humanas en la red global de internet, como: Tuits, llamadas, mensajes de texto, así como los diferentes registros generados a partir de la actividad económica; es decir pagos con medios electrónicos, son datos que pueden llegar a ser utilizados en: la generación de una alerta temprana de cómo la población se está viendo afectada por una situación de desastre adversa en una región específica; así como una herramienta en la planificación de las intervenciones humanitarias, mediante la gestión eficaz de los recursos, permitiendo así centrarse en los lugares que requieran más atención.

(...) La división de análisis de datos de Naciones Unidas, Global Pulse, y la de BBVA, *BBVA Data & Analytics*, pusieron en marcha un proyecto para entender la solidez económica de las personas que se enfrentan a desastres naturales analizando los datos de las operaciones financieras y comprobando cómo se comportan antes y después.

Una de las primeras investigaciones que se realizaron sobre el tema fue la del impacto económico que produjo el huracán Odile en el 2015, en el estado de baja california en México, el proyecto arrojó como resultado que, dentro de los diferentes núcleos familiares, se gastó 50% más de lo habitual (BBVA Analytics), en productos básicos como gasolina y alimentos. En casos como los anteriormente nombrados, el Big Data puede ofrecer claves para una gestión más eficiente de los recursos ante los desastres naturales, así como la atención adecuada a la población afectada.

Recientemente las naciones unidas llevo a cabo el primer congreso en torno al manejo y utilización de datos masivos, para enfocarlos hacia el apoyo al cumplimiento de los objetivos de desarrollo sostenible, midiendo su evolución y alcance mediante indicadores respaldados

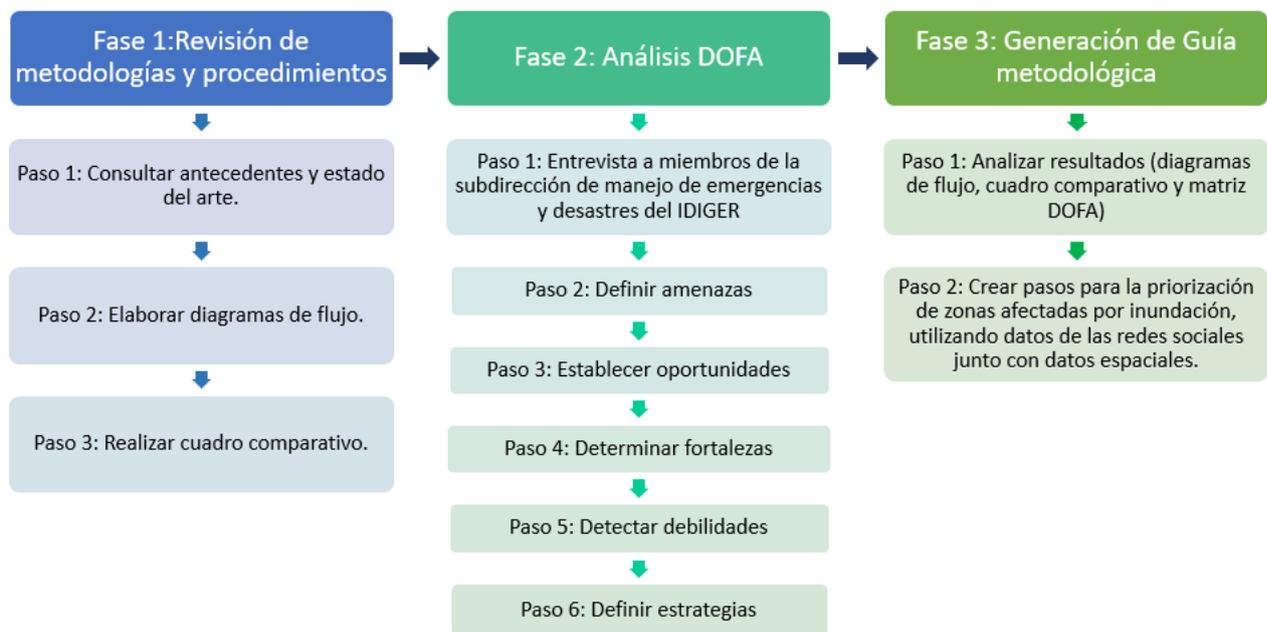
por los mismos datos; razón por la cual la división de análisis de datos de las naciones unidas: ***UNITED NATIONS GLOBAL PULSE: harnessing big data for development and humanitarian action*** ), ha estructurado un conjunto de proyectos encaminados a la utilización del Big Data para resolver los problemas más apremiantes de la sociedad; dichos proyectos son:

1. Incrementar las taxonomías para ayudar a los esfuerzos de gestión de desastres (**crowdsourcing taxonomies to assist disaster management efforts**): el objetivo de este proyecto se enfoca hacia el cumplimiento de los objetivos de desarrollo sostenible, en donde se le hace especial énfasis a la comunicación como principal problema en la gestión de desastres; la solución a este problema radica básicamente en entender mejor lo que dice y como interactúa la gente dentro de las redes sociales; es utilizar el análisis de datos para extraer información relevante a temas prioritarios, cuyo primer paso es crear taxonomías, conjuntos o palabras clave dentro de los mismos motores de búsqueda.
2. Utilizar los datos de las transacciones financieras para medir la resiliencia económica de la población a los desastres naturales (**using financial transaction data to measure economic resilience to natural disasters**): el objetivo de este proyecto se enfoca en la ayuda humanitaria y explora como se pueden analizar los datos provenientes de las transacciones bancarias enfocadas en analizar la resiliencia económica de la población damnificada por desastres naturales, el área de estudio para este proyecto fue igualmente el estado de baja california en México durante el paso del huracán Odile, donde se evaluaron por medio de indicadores el impacto en los medios de subsistencia y las actividades económicas. El proyecto midió transacciones diarias de puntos de venta y retiros de cajeros automáticos a alta resolución geoespacial para obtener una idea de la forma en que las personas se preparan y se recuperan de un desastre.

## 4. METODOLOGÍA

La metodología propuesta en este trabajo de investigación se desarrolló en tres fases: Revisión de metodologías y procedimientos, análisis DOFA y la generación de una guía metodológica que permitirá reproducir los resultados dentro de este proyecto mediante pasos específicos (Ver figura 7). Dichas fases y metodologías se evidencian y se encuentran descritas a continuación:

Figura 7. Marco metodológico



### 4.1 Fase 1: revisión de metodologías y procedimientos

Esta fase consta de una revisión y evaluación de metodologías y procedimientos relacionados con el uso de los datos generados en las redes sociales para atender emergencias por inundación. En esta parte se desarrollaron los siguientes tres pasos:

**Paso 1. Consultar antecedentes y estado del arte:** este procedimiento se realizó a través de una consulta de información, sobre investigaciones hechas en otros países, en la base de datos de Science Direct y en Google Académico. Algunas de las palabras claves utilizadas en los buscadores fueron: Big Data + Social Media+ Floods + Emergency Response. Luego de esta búsqueda se hizo una lectura y revisión minuciosa de los temas, el área de estudio, contexto y

metodologías utilizadas. En este caso se encontraron 6 trabajos de investigación relacionados con el tema de estudio:

El primer trabajo titulado “Inundación a través del lente de la actividad del teléfono móvil” (Pastor-Escuredo et al., 2014), presenta una metodología, donde se estudió la viabilidad del uso de datos de registros de detalles de llamadas (CDR), combinados con otras fuentes de información, para caracterizar las inundaciones que ocurrieron en Tabasco- México en 2009.

La metodología del estudio tiene tres pasos principales que son: la evaluación de la representatividad de los datos, el análisis de eventos basado en datos y la integración de datos. El primer paso permitió estudiar los patrones de comportamiento social dentro de los datos de CDR y evaluar qué tan representativo es el móvil en la población objetivo. En esta parte utilizaron el censo 2010 de Tabasco como fuente fundamental de información para evaluar la representatividad.

Dentro del análisis de eventos basado en datos se han desarrollado métodos de procesamiento personalizados. Para hacer el análisis de CDR y datos de teledetección se ha implementado una biblioteca de herramientas Python para analizar archivos de bases de datos CDR, filtrar los datos de acuerdo a sus coordenadas GPS asociadas, reconstruir trayectorias de desplazamiento, medir descriptores estadísticos y visualizar todo con datos geo-referenciados. Con el fin de analizar las imágenes de sensores remotos e identificar el área inundada han reconstruido un mapa de impacto utilizando imágenes Landsat-7 y han utilizado una línea de procesamiento de imágenes que emplea una morfología matemática y una probabilidad máxima por píxel.

En la integración de datos, se recogieron y se homogeneizaron los datos en un SIG para analizar las señales de la actividad de comunicación y la comparar los datos de CDR con los niveles de lluvia extraídos de los datos del proyecto NASA-TRMM.

El segundo trabajo de investigación: “Explorando el uso de las redes sociales durante la inundación de 2014 en Malasia” (Aisha, Wok, Manaf, & Ismail, 2015), muestra un procedimiento en el que se distribuyeron cuestionarios a las familias, de la costa este de Malasia, directamente afectadas por la inundación. Se tomó una muestra de 507 personas y se midió con escalas de Likert de 5 puntos el uso de las redes sociales, la probabilidad de usar las redes sociales para compartir información, los comportamientos de intercambio de información

y los sentimientos positivos o la satisfacción obtenida al compartir información sobre la inundación. Este estudio se realizó con el fin de dar una guía sobre cómo se pueden utilizar las redes sociales para compartir información crítica durante los periodos de desastre, donde los medios tradicionales pueden fácilmente no estar disponibles.

El tercer trabajo de investigación: “Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres” (Nair, Ramya, & Sivakumar, 2017), contiene un estudio sobre cómo las personas en Chennai utilizaron las redes sociales, especialmente Twitter, en respuesta a una de las más devastadoras inundaciones del país. Los Tweets recopilados se analizaron con algoritmos de aprendizaje automático como Random Forests (Bosques Aleatorios), Naive Bayes y Decision Tree (Árboles de Decisión).

Al comparar el rendimiento de los tres algoritmos se encontró que Random Forests (Bosques Aleatorios) es el mejor para usar durante un desastre. Además se estableció, en la recolección de datos, que durante el periodo de inundación, en Twitter, casi el 70%-75% de los tweets estaban relacionados con la palabra clave Chennai Flood; también, en la clasificación de datos, se encontró que durante el desastre hubo un mayor número de Tweets (620) en las dos primeras categorías que son “Medidas de alivio” y “Necesidad de Ayuda”, después del desastres hubo un mayor número de Tweets (casi 505) en las siguientes 3 categorías que son “Expresar gratitud”, “Quejas” y “Otros”.

Los Tweets publicados durante el periodo de inundación hicieron referencia al nivel del agua en varias áreas, a solicitudes de más voluntarios para empacar kits de supervivencia, a solicitudes de asistencia médica, suministro de alimentos, agua y barcos. En las últimas semanas después del desastre los Tweets hacían referencia a medidas de socorro emprendidas por el gobierno y organizaciones no gubernamentales, a la reconstrucción de Chennai y la orientación sobre el proceso para recibir compensación por parte del gobierno. Este trabajo de igual forma exploró los usuarios de Twitter más influyentes que comentaron sobre la inundación.

En la cuarta metodología desarrollada en la investigación: “Dominar las redes sociales: un análisis de las comunicaciones del condado de Jefferson durante las inundaciones de Colorado 2013” (Hiltz et al., 2013), se indagó sobre las comunicaciones en las redes sociales y las prácticas de trabajo del Equipo de Gestión de Incidentes Tipo III del condado de Jefferson

durante las inundaciones de Colorado en septiembre de 2013 a través de la evaluación de las comunicaciones sobre la inundación en Facebook, Twitter y el blog del equipo para conocer cómo este equipo coordinó sus comunicaciones para satisfacer las necesidades de información de la comunidad. Usando un enfoque de método mixto de entrevistas y análisis de contenido en las redes sociales, se describió el comportamiento del equipo en línea en relación con las necesidades de respuesta de la emergencia.

Además se informó sobre las adaptaciones de las prácticas laborales del equipo para extender las comunicaciones tradicionales con las redes sociales y así crear un plan de comunicación integrado. Los principales procesos que se llevaron a cabo en el desarrollo de esta investigación son: la realización de entrevistas al equipo, la recopilación de datos y su codificación en 8 categorías temáticas como: Hipervínculos y referencias explícitas, compromiso en línea, protocolo, público, rumor, seguridad, servicio y estado con sus respectivas subcategorías.

La quinta metodología utilizada en el trabajo de investigación: Chatter on The Red: ¿Qué revela la amenaza sobre la vida social de la información generada en microblogs? (Starbird, Palen, Hughes, & Vieweg, 2010), considera la comunicación mediada por computadora (CMC) durante la inundación que tuvo lugar en el Valle del Río Rojo en los Estados Unidos y Canadá en marzo y abril de 2009; tiene un enfoque en el uso de Twitter para identificar mecanismos de producción, distribución y organización de la información. El evento de Red River dio lugar a una rápida generación de comunicaciones en Twitter por numerosas fuentes que utilizan una variedad de formas de comunicación como informes autobiográficos o medios de comunicación principales.

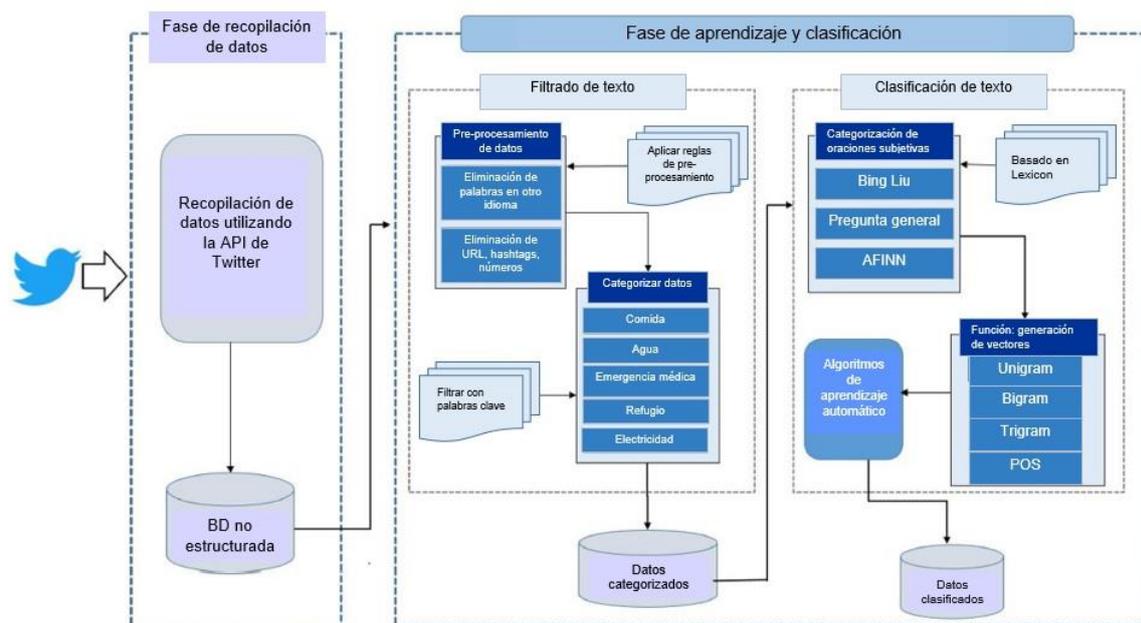
Este estudio examinó la vida social de la información generada en microblogs teniendo en cuenta que el panorama de Twitter es tal que la producción de nueva información se apoya sobre actividades derivadas de dirección, transmisión, síntesis y redistribución; además se complementa con la innovación socio técnica. Es por esto que la investigación se centró en las comunicaciones de Twitter que tuvieron lugar durante un periodo de 51 días del evento de inundación. La ventana de recolección de datos comenzó desde el 8 de marzo de 2009 hasta el 28 de abril y los principales pasos para el desarrollo de esta investigación son: la recolección de datos, codificación cualitativa de datos y el examen de otras fuentes en línea.

La sexta metodología utilizada en la investigación: “Big Data Analytics para la respuesta ante desastres y recuperación a través del análisis de sentimientos” (Ragini, Anand, & Bhaskar, 2018), propone un enfoque basado en Big Data para la respuesta a desastres a través del análisis de sentimientos. El modelo propuesto recopila datos de desastres por inundación de las redes sociales, en varios países del sur de Asia, y los clasifica de acuerdo con las necesidades de las personas afectadas. Los datos de desastres categorizados se clasifican con un algoritmo de aprendizaje automático para analizar el sentimiento de las personas.

Además se analizan diversas características de los tweets como partes del habla y el léxico para identificar la mejor estrategia de clasificación para los datos de desastres (ver figura 8).

En primer lugar se recopilaron datos de las inundaciones en India- Pakistán en septiembre de 2014 (70.817 tweets), de la tormenta ciclónica HUDHUD desde el 05/09/14 al 15/11/14 (30.000 tweets), la inundación en Kashmir (30,817 tweets) y la tormenta ciclónica llamada Nilofar desde el 05/09/14 al 15/11/14 (10.000 Tweets). Las palabras clave utilizadas para recopilar los datos fueron: HUDHUD, Vizag flood, Nilofar, Kashmir floods, India-Pakistan floods y Pakistan floods; utilizando la API de Twitter de transmisión.

Figura 8. Ragini, Anand, & Bhaskar, 2018. Big Data Analytics para la respuesta ante desastres y recuperación a través del análisis de sentimientos.



Recuperado de: <https://doi.org/10.1016/J.IJINFOMGT.2018.05.004>

Luego se evaluó la fase de aprendizaje y clasificación del modelo propuesto a través de dos métodos:

- El primer método consiste en el filtrado de texto donde los datos de desastre se procesan previamente para que el algoritmo de aprendizaje automático en la siguiente etapa puede comprender los datos. En el caso de mensajes de Twitter, los datos de entrada contienen URLs, números, palabras en idiomas extranjeros, abreviaturas, símbolos y emoticones. El pre- procesamiento de datos se realiza para eliminar todos estos caracteres innecesarios de los datos de entrada. Una vez que se realiza el pre- procesamiento de datos, los datos de entrada se clasifican de acuerdo con las diversas necesidades de las personas.

Esta clasificación se hace mediante el filtrado de palabras clave. Las palabras clave se seleccionan identificando palabras que se encuentran más de cinco veces y que son relevantes para cada una de las categorías. Estas palabras aplicadas al conjunto de datos recopilados y filtrados resultaron en 6842 tweets.

Cabe aclarar que, en la lista de categorías, la estructura colapsada y las personas atrapadas no son consideradas. Para ayudar a las personas afectadas, se debe llevar a cabo un tipo diferente de análisis (análisis de estimación de costos para la estructura colapsada y plan de rescate de emergencia para las personas atrapadas). El resto de las categorías incluyen alimentos, agua, refugio, emergencias médicas y electricidad, las cuales son las necesidades más exigentes de las personas afectadas por el desastre.

- El segundo método se refiere a la clasificación de texto, el cual tiene 2 subprocesos: el primero segrega la oración subjetiva al filtrar cada categoría del texto relacionada con los datos del desastre con los léxicos considerados Bing Liu, General Inquirer y AFINN. Se llevan a cabo partes del etiquetado de voz para todas las oraciones subjetivas y se descartan las oraciones objetivas (Neutral). Las oraciones subjetivas se convierten en vector de caracteres mediante la aplicación de adjetivos, adverbios y en caracteres de la bolsa de palabras.

El segundo subproceso segrega las frases subjetivas de todo el flujo de textos relacionados con desastres. Las frases subjetivas se identifican al escanear todo el texto con los léxicos de Bing Liu, General Inquirer y AFINN. Una vez que se identificaron las palabras positivas o negativas, la palabra se toma con un tamaño de ventana de tres. Estas frases subjetivas se convierten en vector de caracteres aplicando las características de unigram, bigram, trigram y basado en partes del discurso (POS). El algoritmo Support Vector Machine (SVM) se entrena con los vectores de características resultantes. La clasificación basada en SVM tiene dos niveles: Nivel de entrenamiento y clasificación. En SVM, el nivel de entrenamiento construye un modelo de entrenamiento que se usa para predecir el texto de entrada como positivo o negativo en el nivel de clasificación.

Los resultados del algoritmo de aprendizaje automático se evaluaron calculando la precisión, recuperación y F-measure. Estas son las tres medidas principales para calcular el desempeño de un sistema de clasificación. Para evaluar estas medidas, la clase bajo investigación se considera positivo y todas las demás clases restantes como negativas. La precisión es una fracción del texto clasificado que es relevante. La recuperación es una fracción del texto clasificado que se recupera. F-measure es la relación de la combinación de precisión y recuperación. El método propuesto se implementa y se prueba utilizando el marco de Big Data de Apache Spark y el lenguaje de programación Python.

Paso 2. Elaborar diagramas de flujo: con el fin de resumir las metodologías empleadas en cada investigación consultada, se realizaron 5 diagramas de flujo con la herramienta online Lucid Chart, que se puede encontrar en la página: (<https://www.lucidchart.com/es-LA/users/login>).

Paso 3. Realizar cuadro comparativo: en este paso se hizo un cuadro comparativo que resume el contexto, los objetivos, los medios o las herramientas utilizadas y los pasos metodológicos de cada trabajo de investigación. Es necesario aclarar que en el contraste de las 6 investigaciones relacionadas con el tema de estudio, hay varias similitudes y diferencias que permiten identificar la metodología más conveniente para usar los datos Big Data en la atención a emergencias.

## 4.2 Fase 2: análisis DOFA

La fase 2 consta de un análisis de los procesos de la subdirección del IDIGER, frente a la atención de desastres por inundaciones. Esta se desarrolló mediante la recolección de información primaria dentro de la entidad y determinación de la forma más pertinente para hacer el análisis, que fue la construcción de una matriz DOFA, la cual permitió definir los aspectos internos (Fortalezas y Debilidades) y los externos (Oportunidades y Amenazas) que en cierta medida influyen en el favorecimiento o involución de la entidad.

Para las fortalezas se tuvo en cuenta las características de la entidad que la hacen superior o le dan un valor agregado frente otras entidades en cuanto a la atención y manejo de inundaciones; para las debilidades se tuvo en cuenta los problemas y las limitaciones que impiden a la entidad lograr una adecuada atención, las oportunidades se evaluaron frente a cada factor del entorno de la entidad que pueden proporcionar aspectos positivos y que se pueden aprovechar; y finalmente para las amenazas se evaluaron todas aquellas situaciones que pueden obstaculizar o impedir que se alcancen los objetivos de la entidad en cuanto la priorización de atención a desastres por inundaciones. Lo anterior es una razón por la cual se establecieron 6 pasos para su elaboración y evaluación dentro de los procesos de la subdirección de manejo de emergencias y desastres del IDIGER:

Paso 1. Entrevista a miembros de la subdirección de manejo de emergencias y desastres del IDIGER: en este paso se hizo una recolección primaria de información dentro de la entidad, es decir que se realizaron visitas a la subdirección de la entidad, ubicada en la Cr127 #22g-18, interior 1 y 2, agrupación industrial La Esperanza, y se llevaron a cabo dos reuniones con dos de los encargados del equipo de trabajo subdirección para el manejo de emergencias y desastres: Claudia Patricia Coca Galeano y Jorge Suárez Stevenson.

Paso 2. Definir amenazas: se determinaron los factores que suponen riesgos para la entidad debido a que son factores externos que no tienen dependencia directa del IDIGER por lo que son difícilmente controlables, para determinar las amenazas de la entidad frente a la atención a emergencias por inundaciones se evaluaron aspectos específicos como: ¿Cuáles son las necesidades de la población en una situación adversa?, ¿alguna debilidad establecida puede significar una amenaza para la entidad?, ¿Cómo están afectando a la entidad las medidas del

gobierno?, ¿qué obstáculos puede o está enfrentado la entidad?, y así mismo se priorizaron las amenazas según la severidad y probabilidad de ocurrencia.

Paso 3. Establecer oportunidades: al igual que las amenazas, las oportunidades son factores externos a la entidad, es decir que son difícilmente controlables al no existir control directo, pero si se pueden desarrollar planes o estrategias para aprovecharlas, ya que se evaluaron las posibles ventajas que la entidad puede aprovechar y mejorar de manera gradual, lo que resulta en una ventaja comparativa y competitiva, razón por la cual se evaluaron aspectos como: las tendencias globales y nacionales, cambios tecnológicos que pueden representar una oportunidad, como se comportan las personas ante determinada situación, comparación de entidades entre otras.

Paso 4. Determinar fortalezas: a diferencia de los dos anteriores, las fortalezas corresponden a los factores internos de la entidad y que representan aspectos positivos, para este caso se evaluaron las fortalezas al nivel de procesos, para determinar las fortalezas de la entidad se tomaron en análisis aspectos como: qué los hace mejores frente a otras entidades, cuál es el valor agregado de la entidad, efectividad dentro de los procesos y los correspondientes puntos fuertes dentro de cada proceso implicado en la atención a emergencias.

Paso 5. Detectar debilidades: se evaluaron todos aquellos aspectos que, aunque están bajo control al ser un aspecto a evaluar interno de la entidad, pueden significar una desventaja para la atención a emergencias por inundaciones, aspectos como: en qué se puede mejorar, qué se debe dejar de hacer en aras de mejorar, en cuales procesos falta experiencia u organización, entre otros.

Paso 6. Definir estrategias: se trazaron comparando las fortalezas, oportunidades, amenazas y debilidades, es decir, comparar de manera transversal los factores internos y externos de la entidad, de forma que se puedan trazar estrategias que permitan usar las fortalezas para aprovechar las oportunidades, así mismo se evaluó como se pueden enfocar las fortalezas para mitigar las amenazas, y las oportunidades para mejorar, eliminar o corregir las debilidades, lo anterior se ve reflejado mediante un arreglo matricial en los resultados.

### 4.3 Fase 3: generación de guía metodológica

Esta fase plantea el componente propositivo de esta investigación, en donde se propondrá una metodología que indique los pasos a seguir y las herramientas necesarias para usar los datos de las redes sociales en los procesos de atención a desastres por inundación. Para la generación de tal guía se siguieron los siguientes dos pasos:

Paso 1. Analizar resultados: a partir de los resultados obtenidos como diagramas de flujo, cuadro comparativo y análisis DOFA se estableció la mejor metodología para basar la propuesta de este trabajo de investigación y esto se relacionó con el análisis DOFA de la subdirección para dar una propuesta que contribuya a la mejora de las debilidades de la entidad.

Paso 2. Crear pasos para la priorización de zonas afectadas por inundación, utilizando datos de las redes sociales junto con datos espaciales: para la creación de la primera parte de la propuesta de uso de datos de las redes sociales para la atención de emergencias causadas por inundaciones, se consideró el estudio realizado en Chennai denominado: Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres (Nair, Ramya, & Sivakumar, 2017). La metodología de este trabajo se resume en una recolección, filtrado y clasificación de Tweets con determinadas herramientas.

La recolección de datos de Twitter se construyó consultando a un ingeniero de sistemas sobre los métodos y elementos necesarios para tal fin. Se hizo un ejemplo utilizando cuentas de Twitter de entidades oficiales, noticias y hashtags que se usan actualmente como #TemporadaDeLluvias, #ÁrbolEnRiesgo y #Río Bogotá y se corrió un código Java a través de la herramienta Eclipse.

En cuanto al filtrado y clasificación de los Tweets se tuvo en cuenta la herramienta Weka, utilizada en el trabajo de investigación revisado, se consultaron varios manuales y tutoriales para conocer el entorno de trabajo, funcionalidades y métodos para procesar los datos. A partir de esto se sugirió un procedimiento general que sirve para procesar los Tweets relacionados con inundaciones, así como otros nombres para los hashtags y etiquetas de clasificación.

La segunda parte de la propuesta, entendida como el uso de datos espaciales y herramientas geográficas a partir de los datos generados en las redes sociales, tiene como un paso

metodológico esencial la creación de niveles de atención a damnificados y/o afectados, en donde se priorizan los lugares con mayor demanda de atención haciendo lo siguiente:

1. Determinar del tipo de datos y fuentes (ver anexo 5) a tener en cuenta para encaminar la priorización de acuerdo a las diferentes necesidades de las personas frente a una situación adversa, apoyado en la minería de datos.
2. Establecer la ubicación de las zonas con mayor volumen de datos, mediante el uso de la georreferenciación implícita en cada dato que se va a utilizar, pues hoy en día cualquier dato este, o se puede georreferenciar.
3. Generar zonas de influencia teniendo como punto central, los puntos con mayor volumen de datos, así como teniendo en cuenta el tipo y la magnitud del desastre, mediante la utilización de herramientas SIG como lo son polígonos de Thiessen e interpolación IDW.
4. Extrapolar y estimar el número de personas afectadas teniendo en cuenta los “*Hot Spot*” o puntos calientes de información proveniente de las fuentes seleccionadas anteriormente.
5. Producir un modelo/simulación, en el cual se puedan ver en acción los diferentes niveles de atención y se evidencie su pertinencia como modelo mediante la creación de cartografía que represente de forma gráfica las zonas afectadas haciendo más fácil y posible su priorización.

## 5. RESULTADOS Y ANÁLISIS DE RESULTADOS

Los resultados obtenidos en este trabajo de investigación se presentan en tres partes importantes que son: Una comparación de metodologías de aplicación de Big Data en otros lugares del mundo, un análisis de debilidades, oportunidades, fortalezas y amenazas del IDIGER y finalmente una guía metodológica para priorizar la atención a emergencias generadas por inundaciones en una zona específica de la ciudad de Bogotá.

### 5.1 Comparación de metodologías de aplicación del Big Data en otros países

A continuación, se presentan los resultados obtenidos del primer objetivo. Es importante destacar que luego de la revisión de diversas investigaciones, solo 6 coinciden con el tema trabajado en esta investigación. En las figuras que se encuentran luego de esta descripción, se pueden observar los diagramas de flujo correspondientes a las metodologías de aplicación del Big Data en la atención a emergencias por inundaciones en diferentes lugares.

Es necesario señalar que los diagramas de flujo son representaciones gráficas que explican un proceso mediante figuras que indican el flujo y un orden claro de los procedimientos realizados. Las figuras 9, 10, 11, 12 y 13 muestran los diagramas de flujo para cada una de las metodologías usadas en los proyectos de investigación consultados y relacionados con el uso de los datos de las redes sociales en la atención a emergencias por inundaciones. Todos los diagramas tienen figuras que se pueden entender de la siguiente manera: El óvalo naranja representa el inicio de un proceso, el rombo amarillo indica una decisión, los cuadrados azules son los procesos principales y los cuadrados verdes los subprocesos dentro de los procesos principales.

La figura 9 muestra el trabajo de investigación sobre la viabilidad de usar datos de registros de detalles de llamadas (CDR), combinados con otras fuentes de información, para caracterizar el impacto de las inundaciones en Tabasco- México 2009. Los procesos principales de este trabajo se resumen en una evaluación de la representatividad de los datos, análisis de eventos basado en datos e integración de datos; en cada proceso principal se observa que para solucionar la pregunta inicial, se utilizan herramientas y datos cartográficos que complementan el uso de datos de llamadas.

La figura 10 presenta la metodología utilizada en la investigación referente al uso de las redes sociales durante la inundación de Malasia 2014. Los procesos principales que se desarrollaron fueron: la distribución de encuestas y mediciones del uso de las redes sociales, de la probabilidad de uso, de los comportamientos al compartir información y de los sentimientos positivos; dentro de los sub procesos realizados se evidencia que solamente se emplearon técnicas estadísticas para medir el uso de las redes sociales de una muestra de una población.

La figura 11 enseña los procedimientos realizados en el trabajo que estudia la forma en que las personas en Chennai utilizaron Twitter en respuesta a la inundación del año 2015. Los procesos principales que se llevaron a cabo se definen en: la recolección de Tweets en una ventana de tiempo específica, el procesamiento de datos y la clasificación de los tweets. Dentro de estos procesos hay herramientas y métodos más específicos que permiten deducir un mayor uso del Big Data con técnicas y herramientas especiales.

La figura 12 muestra los métodos usados en la investigación donde se indagó acerca de las comunicaciones en las redes sociales y las prácticas de trabajo del Equipo de Gestión de Incidentes Tipo III del condado de Jefferson durante las inundaciones de Colorado en septiembre de 2013. Los pasos primordiales desarrollados en este trabajo fueron: la elaboración de entrevistas al equipo, la recopilación de datos y codificación. Lo anterior refleja que este trabajo integra las entrevistas a una entidad local encargada de atender emergencias y el uso de los datos de las redes sociales para potenciar el desarrollo de las actividades de dicha entidad.

La figura 13 representa los procedimientos ejecutados en la investigación que considera la comunicación mediada por computadora (CMC) durante la inundación que tuvo lugar en el Valle del Río Rojo en 2009; esta tiene un enfoque en el uso de Twitter para identificar mecanismos de producción, distribución y organización de la información. Las fases primordiales que hacen parte de esta metodología son: la recolección de datos, la codificación y la evaluación de otras fuentes en línea, esto permite inferir en los subpasos, un uso del Big Data con técnicas de visualización que no están relacionadas con el uso de técnicas cartográficas.

Finalmente, la tabla 3 muestra la comparación de los proyectos de investigación descritos anteriormente, contiene un resumen puntual del contexto, objetivos, herramientas empleadas

(medios) y los pasos ejecutados en cada trabajo. Así mismo presenta las similitudes y diferencias observadas durante la revisión.

Figura 9. Metodología 1: Inundación a través del lente de la actividad del teléfono móvil

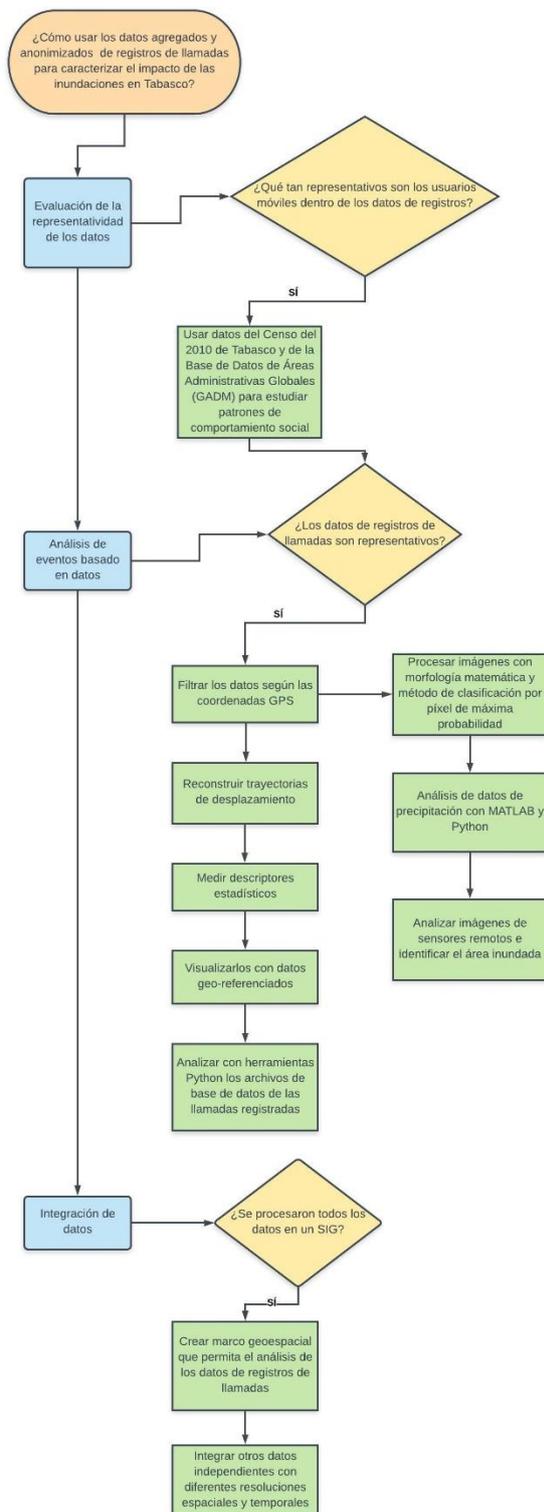


Figura 10. Metodología 2: Explorando el uso de las redes sociales durante la inundación de 2014 en Malasia

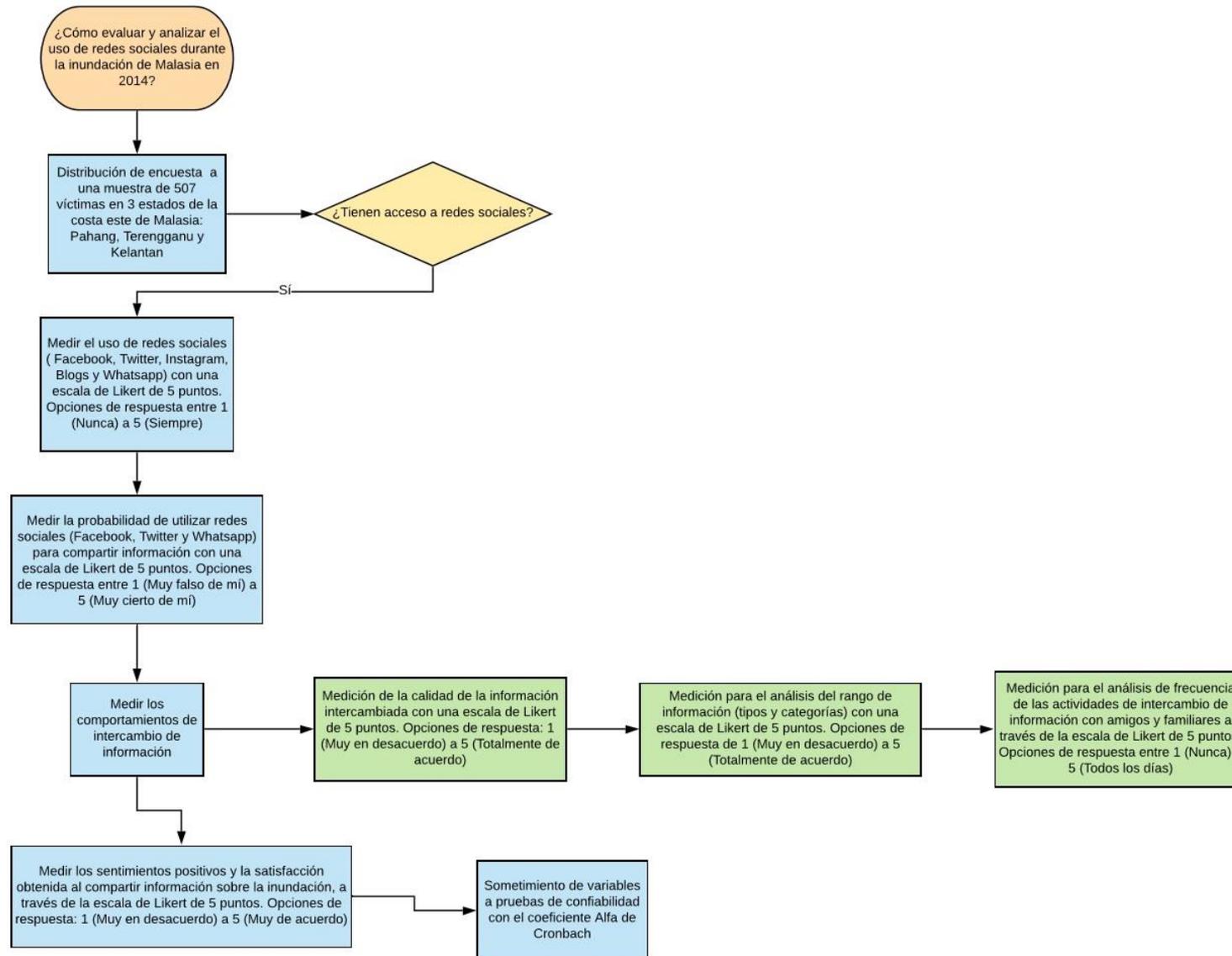


Figura 11. Metodología 3: Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres

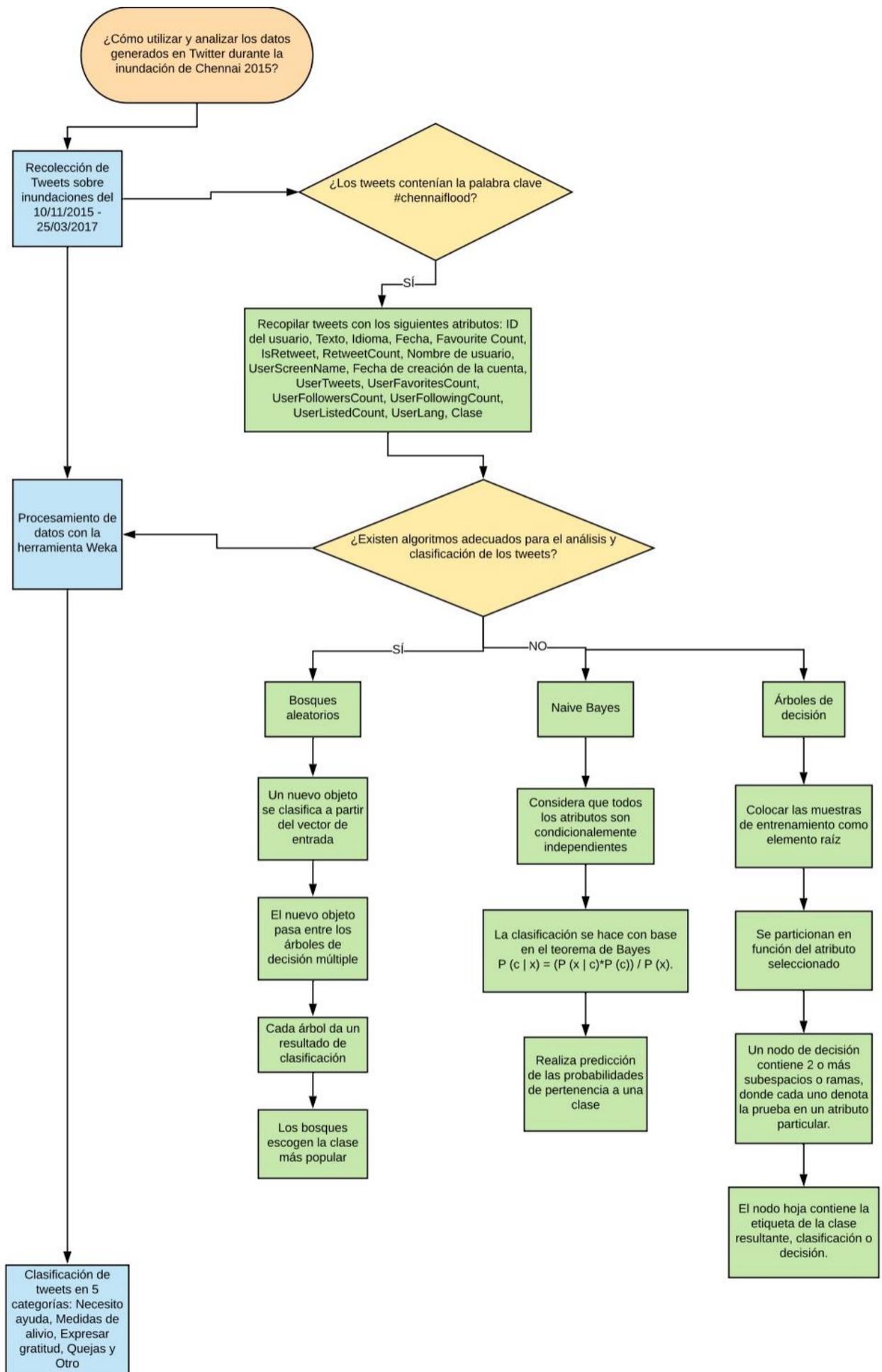


Figura 12. Metodología 4: Dominar las redes sociales: un análisis de las comunicaciones del condado de Jefferson durante las inundaciones de Colorado 2013

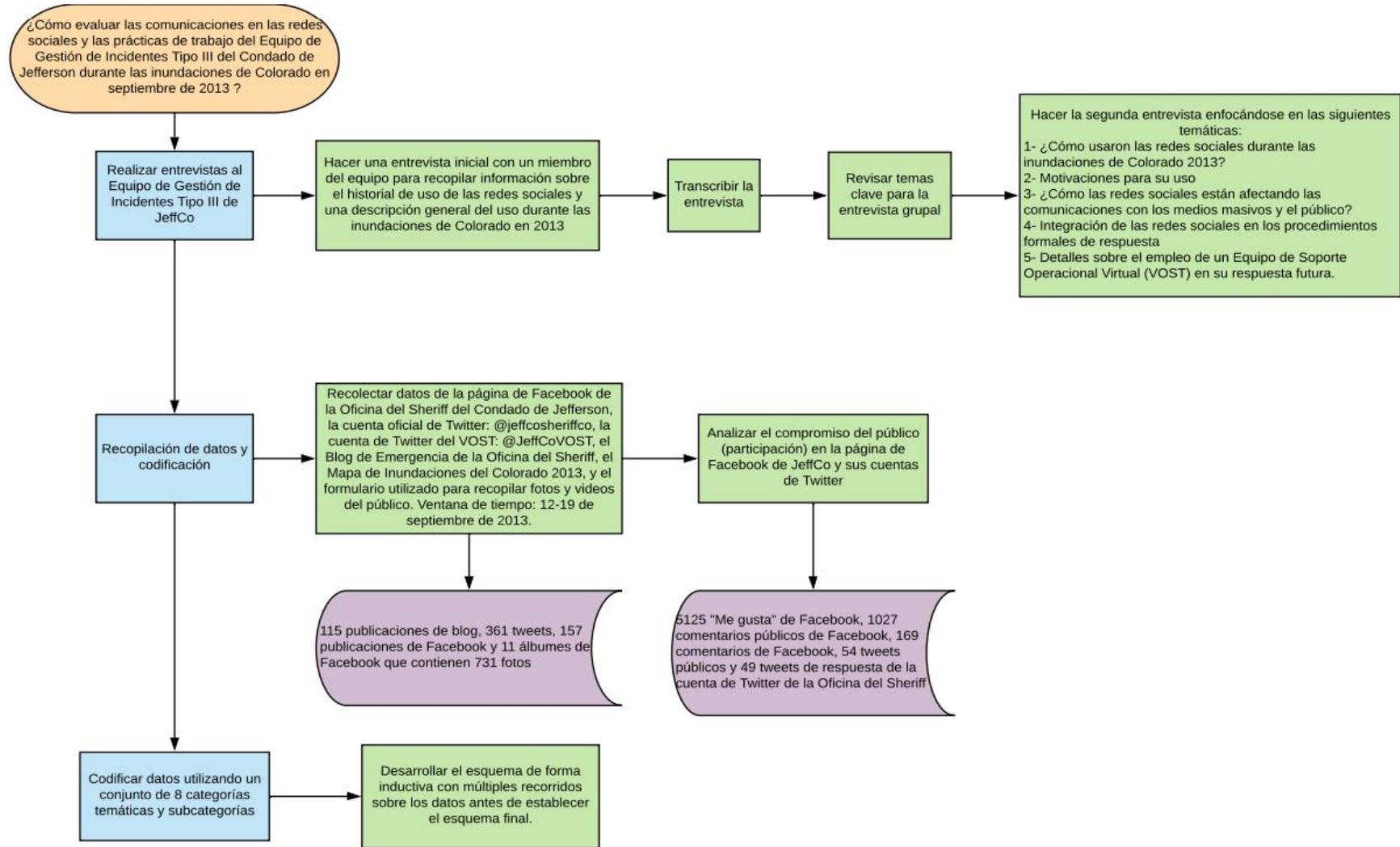


Figura 13. Metodología 5: Chatter on The Red: ¿Qué revela la amenaza sobre la vida social de la información generada en microblogs?

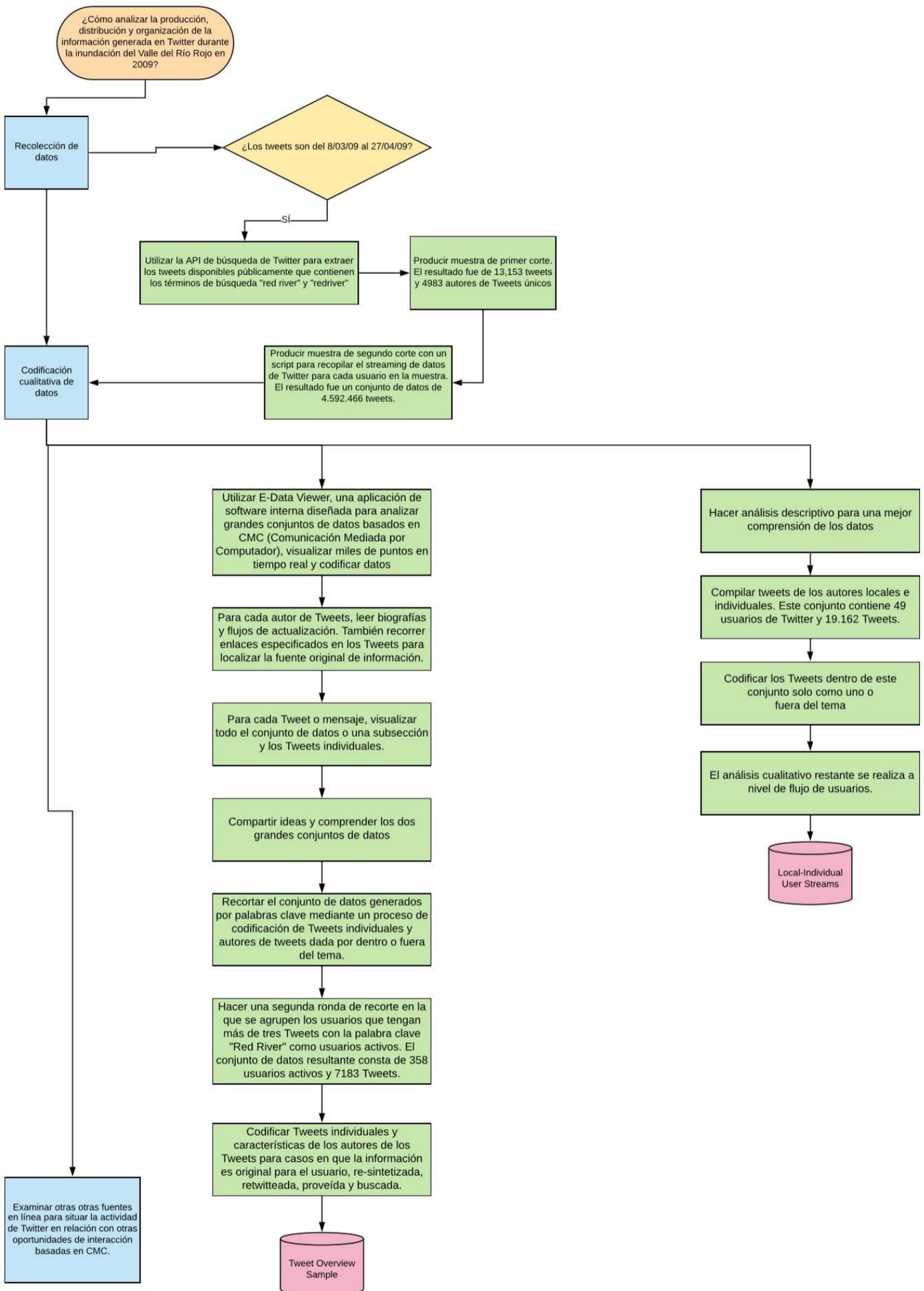


Tabla 3. Comparación: modelos metodológicos de aplicación Big Data.

<b>CUADRO COMPARATIVO</b>				
<b>PROYECTO</b>	<b>RESUMEN</b>	<b>OBJETIVOS</b>	<b>HERRAMIENTAS EMPLEADAS</b>	<b>PASOS</b>
<b>Inundación a través del lente de la actividad del teléfono móvil</b>	Este trabajo presenta una indagación de la viabilidad de usar datos de CDR (Call Detail Records) combinados con otras fuentes de información para caracterizar las inundaciones que ocurrieron en Tabasco, México en 2009.	<ul style="list-style-type: none"> <li>- Desarrollar y aplicar métodos para evaluar la idoneidad del uso de datos CDR agregados y anónimos para caracterizar el impacto de las inundaciones en las poblaciones, utilizando las inundaciones de Tabasco, México en 2009.</li> <li>- Contribuir al desarrollo de herramientas de apoyo a la toma de decisiones basadas en CDR en tiempo real para la respuesta del sector público ante inundaciones y otros desastres naturales.</li> </ul>	<ul style="list-style-type: none"> <li>- Datos de registros de detalles de llamadas (CDR).</li> <li>- Fuentes de datos adicionales: Imágenes satelitales, datos de precipitación, datos de protección civil, datos del censo, datos de áreas administrativas.</li> <li>- Otra información contextual.</li> <li>- SIG.</li> <li>- MATLAB.</li> <li>- Python.</li> </ul>	<ul style="list-style-type: none"> <li>- Recolección de datos.</li> <li>- Evaluación de la representatividad de los datos: evaluar representatividad de los usuarios dentro de los registros.</li> <li>- Análisis de eventos basado en datos: procesamiento de datos geoespaciales.</li> <li>- Integración de datos: integrar otros datos con diferentes resoluciones espaciales y temporales.</li> </ul>
<b>Explorando el uso de las redes sociales durante la inundación de 2014 en Malasia</b>	Esta investigación muestra una exploración del uso de las redes sociales para compartir información sobre la inundación de Malasia en Diciembre de 2014	<ul style="list-style-type: none"> <li>- Explorar el uso de las redes sociales durante la inundación de 2014.</li> <li>- Examinar el uso de las redes sociales para compartir información durante la inundación de 2014</li> <li>- Analizar los efectos del uso de las redes sociales para compartir información sobre los comportamientos de intercambio de información y la satisfacción obtenida al compartir información durante la inundación.</li> </ul>	<ul style="list-style-type: none"> <li>- Encuestas: distribuidas a una muestra de 507 víctimas en tres estados de la costa este de Malasia: Pahang, Terengganu y Kelantan.</li> <li>- Escala de Likert.</li> <li>- Operaciones estadísticas.</li> </ul>	<ul style="list-style-type: none"> <li>- Diseño de la encuesta.</li> <li>- Distribución de la encuesta.</li> <li>- Mediciones.</li> <li>- Análisis.</li> </ul>

<p><b>Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres</b></p>	<p>Este estudio presenta un análisis de la forma en que la gente de Chennai usaba las redes sociales, especialmente twitter, durante la inundación del año 2015.</p>	<ul style="list-style-type: none"> <li>- Analizar los Tweets recolectados utilizando algoritmos de aprendizaje automático como Random Forests, Naive Bayes y Decision Tree.</li> <li>- Comparar el rendimiento de los tres algoritmos durante un desastre.</li> </ul>	<ul style="list-style-type: none"> <li>- Tweets.</li> <li>- Weka, es una herramienta analítica de minería de datos gratuita, dentro de esta se clasificaron los Tweets con algoritmos de aprendizaje automático como: Decision trees, Random Forests y Naive Bayes.</li> </ul>	<ul style="list-style-type: none"> <li>- Recolección de datos.</li> <li>- Procesamiento de datos.</li> <li>- Clasificación en 5 categorías.</li> <li>- Análisis de los Tweets.</li> <li>- Evaluación del desempeño y comparación.</li> </ul>
<p><b>Dominar las redes sociales: un análisis de las comunicaciones del condado de Jefferson durante las inundaciones de Colorado 2013</b></p>	<p>Este estudio explica y evalúa la relación entre las comunicaciones en las redes sociales y las prácticas de trabajo del Equipo de Gestión de Incidentes Tipo III del Condado de Jefferson durante las inundaciones de Colorado en septiembre de 2013.</p>	<ul style="list-style-type: none"> <li>- Examinar las comunicaciones relacionadas con las inundaciones en tres plataformas: Facebook, Twitter y el blog del equipo para conocer la coordinación de sus comunicaciones para satisfacer las necesidades de información de la comunidad.</li> <li>- Describir los comportamientos del equipo en línea en relación con las necesidades de la respuesta de emergencia en su conjunto.</li> <li>- Informar sobre las adaptaciones que hay al extender las comunicaciones tradicionales con las redes sociales para crear un plan de comunicación integrado.</li> </ul>	<ul style="list-style-type: none"> <li>- Entrevista inicial y grupal al Equipo de Gestión de Incidentes tipo III del condado de Jefferson.</li> <li>- Facebook.</li> <li>- Twitter.</li> <li>- Blog de emergencia.</li> </ul>	<ul style="list-style-type: none"> <li>- Diseño de la entrevista inicial y grupal.</li> <li>- Aplicación de entrevistas.</li> <li>- Recolección de datos.</li> <li>- Análisis.</li> <li>- Codificación de datos en 8 categorías.</li> </ul>

<p><b>Chatter on The Red: ¿Qué revela la amenaza sobre la vida social de la información generada en microblogs?</b></p>	<p>Esta investigación presenta la identificación de maneras de producción, distribución y organización de la información, considerando las comunicaciones generadas en Twitter durante la inundación del Valle del Río Rojo en EE.UU. y Canadá en marzo y abril de 2009.</p>	<p>-Teorizar sobre el chat basado en CMC en la nueva era de las aplicaciones y servicios de "redes sociales" o Web 2.0. - Describir las características de la relación que tiene el chat con los eventos de emergencia masivos para poder predecir con mayor precisión su potencial en un futuro.</p>	<ul style="list-style-type: none"> <li>- API de búsqueda de Twitter.</li> <li>- E- Data Viewer.</li> <li>- Otras fuentes en línea.</li> </ul>	<ul style="list-style-type: none"> <li>- Recolección de datos.</li> <li>- Producción de muestras: filtro de datos.</li> <li>- Codificación cualitativa de datos.</li> <li>- Análisis.</li> <li>- Evaluación de otras fuentes en línea.</li> </ul>
<p><b>Big Data Analytics para la respuesta ante desastres y recuperación a través del análisis de sentimientos.</b></p>	<p>Este trabajo propone un modelo, con un enfoque basado en Big Data, que recopila los datos de las inundaciones en India-Pakistán en septiembre de 2014 y los clasifica de acuerdo con las necesidades de los afectados.</p>	<ul style="list-style-type: none"> <li>- Analizar y categorizar las diversas necesidades de las personas durante y después del desastre.</li> <li>- Identificar varias características como: bolsa de palabras, partes de características basadas en el tipo de voz y en la lógica; y el mejor algoritmo de rendimiento para cada categoría.</li> <li>- Proponer un método que visualice los sentimientos sobre las necesidades básicas y ayude a mejorar los servicios de emergencia.</li> </ul>	<ul style="list-style-type: none"> <li>- API de Twitter.</li> <li>- Lexicon: Bing Liu, Pregunta General y AFINN</li> <li>- Algoritmo SVM.</li> <li>- Apache Spark.</li> </ul>	<ul style="list-style-type: none"> <li>- Recopilación de datos.</li> <li>- Filtrado de texto: pre-procesamiento de datos.</li> <li>- Clasificación de texto: categorización de oraciones subjetivas.</li> <li>- Implementación del modelo.</li> </ul>

## 5.2 Análisis DOFA de la subdirección de manejo de emergencias y desastres del IDIGER.

A continuación, se evidencia la matriz DOFA, como resultado del análisis realizado a los procesos dentro de la subdirección de manejo de emergencias y desastres del IDIGER. La matriz DOFA es una herramienta que es comúnmente utilizada para la formulación y evaluación de estrategias, que generalmente son utilizadas por empresas y entidades; su nombre proviene de las siglas de las palabras: Debilidades, Oportunidades, Fortalezas y Amenazas.

Las fortalezas (F) y las debilidades (D), son factores internos, que pueden agregar o quitar valor, se incluyen los activos, las habilidades, los recursos y los valores agregados. Las oportunidades (O) y las amenazas (A), son factores externos, por lo que están fuera del control de la entidad, se incluyen factores como la demografía, el comportamiento de las personas, las demás entidades, factores legales y económicos, entre otros.

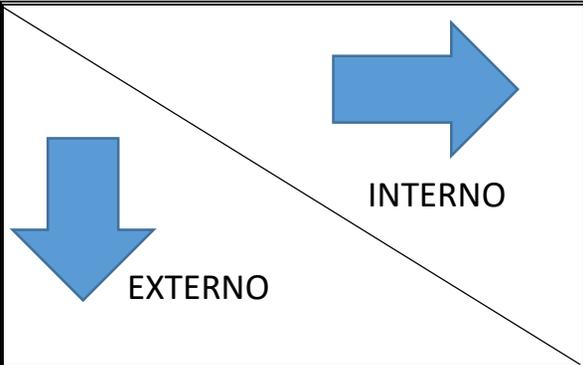
La siguiente matriz DOFA representa un diagnóstico real de la subdirección de manejo de emergencias y desastres del IDIGER, que evidencia la situación real y actual de la misma, y por lo tanto es una herramienta requerida en el proceso de toma de decisiones, teniendo en cuenta también la integración de la información obtenida de los datos generados en las redes sociales dentro de los procesos que apoyen la adecuada gestión dentro de las respuestas institucionales, enfocadas a la atención de emergencias causadas por inundaciones; teniendo en cuenta la revisión de las diferentes metodologías y procedimientos sobre el uso de las redes sociales, frente al manejo de inundaciones en otras ciudades con condiciones similares a las de Bogotá; como es el caso de Chennai, también conocido como Madrás capital del estado de Tamil Nadu en India, donde se tuvo en cuenta el uso y análisis de Twitter durante la Inundación de 2015 para la gestión de desastres.

Lo anterior, pone en evidencia la importancia de aprovechar de la información obtenida de los datos generados en las redes sociales apoyada con el uso e integración de herramientas cartográficas, por parte de la subdirección de manejo de emergencias y desastres, para priorizar la atención a emergencias generadas por inundaciones, lo cual se puede interpretar como una estrategia transversal dentro de la matriz DOFA, que así mismo, genere un valor agregado dentro

de los procesos de la entidad, es decir, las respuestas institucionales frente a la atención a emergencias por inundaciones.

Con el fin de analizar los procesos dentro de la subdirección de manejo de emergencias y desastres del IDIGER, la matriz DOFA (ver tabla 4) es un método de planificación que permite tener claros los enfoques en cuanto a los aspectos positivos y negativos de la entidad, permitiendo evaluar y analizar los diferentes aspectos negativos internos y amenazas externas, confrontándolos juntos con los aspectos positivos internos y oportunidades externas, con el fin de encontrar estrategias que contribuyan a una mejoría progresiva; para este caso dicha matriz será utilizada como una herramienta de control, debido al tiempo de funcionamiento que lleva la entidad, y permitirá evaluar las posibles fallas que no se han identificado con claridad representando una ventaja para la entidad en el camino hacia la mejora de la misma.

Tabla 4. Matriz DOFA de la subdirección de manejo de emergencias y desastres del IDIGER.

	<p style="text-align: center;"><b>FORTALEZAS</b></p> <ol style="list-style-type: none"> <li>1. El nivel de recepción y conexión que tiene el IDIGER frente a otras entidades.</li> <li>2. Capacidad de coordinación de servicios de emergencia.</li> <li>3. Capacidad para coordinar acciones o ejecutarlas frente a emergencias por inundaciones.</li> <li>4. Integración de otras entidades de carácter mixto: Servicios público-privados.</li> <li>5. Personal capacitado para realizar labor en campo; en el lugar de la emergencia.</li> <li>6. Escalabilidad de la emergencia.</li> </ol>	<p style="text-align: center;"><b>DEBILIDADES</b></p> <ol style="list-style-type: none"> <li>1. Muchos intermediarios dentro del proceso de atención.</li> <li>2. No existe una vía de atención directa.</li> <li>3. Manejo de las bases de datos y formularios de forma física.</li> <li>4. Falta de estandarización de las bases de datos.</li> <li>5. Ausencia de variable geográfica en las bases de datos.</li> <li>6. Desaprovechamiento de la información generada por usuarios de internet en tiempo real.</li> </ol>
<p style="text-align: center;"><b>OPORTUNIDADES</b></p> <ol style="list-style-type: none"> <li>1. Generación de cartografía.</li> <li>2. Georreferenciación de las casas y zonas afectadas por inundaciones.</li> <li>3. Fortalecimiento y/o creación de sistemas de monitoreo electrónicos y automatizados.</li> <li>4. Uso de base de datos para el registro y control de recursos.</li> <li>5. Articulación de los consejos locales.</li> <li>6. Aprovechamiento de la información de las redes sociales: Twitter</li> </ol>	<p style="text-align: center;"><b>ESTRATEGIAS F.O</b></p> <ol style="list-style-type: none"> <li>1. Obtención de información e insumos de diferentes entidades para alimentar la cartografía.</li> <li>2. Generación de rutas de atención efectivas mediante el aprovechamiento de la georreferenciación.</li> <li>3. Reducción de los tiempos de capacitación a personal frente a una emergencia.</li> <li>4. Capacitar personal enfocado al área de las TICs.</li> </ol>	<p style="text-align: center;"><b>ESTRATEGIAS D.O</b></p> <ol style="list-style-type: none"> <li>1. Creación de una línea de atención directa.</li> <li>2. Creación y mantenimiento de mapas de calor de focos de inundaciones en Bogotá.</li> <li>3. Elaboración y realización de bases de datos electrónicas.</li> <li>4. estandarización de las bases de datos.</li> <li>5. Integración de los componentes X e Y a las bases de datos.</li> <li>6. Aprovechamiento de la información obtenida de los datos generados en las redes sociales.</li> </ol>
<p style="text-align: center;"><b>AMENAZAS</b></p> <ol style="list-style-type: none"> <li>1. Las entidades que conforman y ejecutan los servicios de respuesta emergencias son independientes.</li> <li>2. Recursos.</li> <li>3. Capacitación inadecuada de personal externo frente a la atención y manejo de la emergencia: inundaciones.</li> <li>4. Participación de las entidades de acuerdo a las particularidades de la emergencia.</li> <li>5. Participación social y comunitaria en la gestión de riesgos y cambio climático, mediante medios electrónicos.</li> </ol>	<p style="text-align: center;"><b>ESTRATEGIAS F.A</b></p> <ol style="list-style-type: none"> <li>1. Incrementar y mejorar la capacidad de organización y coordinación entre las entidades y el IDIGER.</li> <li>2. Generación de formatos digitales amigables con el usuario de fácil llenado.</li> <li>3. Capacitación de la comunidad para incrementar la participación activa de la misma en redes sociales.</li> <li>4. Procurar que los recursos se agoten en primera instancia de Créditos contingentes y pólizas de seguro.</li> </ol>	<p style="text-align: center;"><b>ESTRATEGIAS D.A</b></p> <ol style="list-style-type: none"> <li>1. Reducción de los tiempos de atención por entidades.</li> <li>2. Capacitación del personal frente al llenado de bases y manejo de medios electrónicos y virtuales.</li> <li>3. Mantener informada y capacitada a la comunidad frente a los nuevos recursos virtuales y medios electrónicos.</li> <li>4. Aprovechamiento de la información generada en las redes sociales.</li> </ol>

Como se nombró anteriormente, la matriz DOFA permitió definir los aspectos internos (Fortalezas y Debilidades) y los externos (Oportunidades y Amenazas) que en cierta medida influyen en el favorecimiento o involución de la entidad, la matriz se subdivide de la siguiente manera:

Debilidades:

- Muchos intermediarios dentro del proceso de atención.

Mediante la revisión y análisis de la información primaria recolectada en las diferentes visitas a las dependencias del IDIGER, se logró determinar que dentro del proceso de atención de desastres, el IDIGER tiene dos funciones específicas: la primera es: coordinar a los diferentes servicios de respuesta con los que cuenta la ciudad frente a emergencias: inundaciones, y la segunda función es: la de ejecutar, es decir, que la entidad también debe encargarse de prestar servicios de emergencia y por lo tanto tiene funciones como dicho y como entidad encargada; sin embargo en todo este proceso existen muchos intermediarios ya que la mayoría de la información y los llamados de emergencia se reciben por medio de un número único nacional de atención al ciudadano: NUSE o mejor conocido como la línea 123; esta es la línea que reúne todos los números de seguridad y emergencias: Policía, Bomberos, Movilidad, Gestión de Riesgos y la Secretaría de Salud y demás servicios de emergencia del distrito en uno solo y que tiene como objetivo garantizar una respuesta adecuada, efectiva, eficaz, oportuna y organizada frente a una situación de emergencia, sin embargo para el IDIGER como entidad distrital encargada de la gestión del riesgo de desastres esto representa una debilidad pues no toda la información sobre las emergencias es reportada, lo que impide que sean alimentadas las bases de datos del mismo; y por lo tanto no apoya el proceso de creación de cartografía temática con la información pertinente a Riesgos por inundaciones en la ciudad de Bogotá, que pueda ser de gran ayuda para las demás entidades y para el mismo IDIGER.

Para el IDIGER reducir los intermediarios en el proceso de atención representa generar un trato más cercano, y conocer de primera mano las necesidades de las personas en una situación de emergencia y ahorrar tiempo en el camino, lo anterior gracias a la aplicación del

Big data y herramientas virtuales; ya que desde la aparición del internet y la era de la tecnología, se ha venido presentando la tendencia de muchos sectores tanto públicos como privados de trasladar la mayoría de su oferta y servicios a internet; una de los motivos principales, es que los intermediarios no siempre general un valor agregado, sino que por el contrario general un sobrecosto o una demora en el proceso, para este caso en el proceso de atención.

- No existe una vía de atención directa.

Retomando el inciso anterior, el IDIGER no cuenta con una línea de atención directa, sino que, depende del número único de seguridad y emergencias (NUSE), mejor conocido como la línea 123, la cual es una plataforma digital que cuenta con una década desde su implementación en el año 2007 (Lozada, 2017), que lamentablemente tiene atraso y no cuenta con un sistema eficiente de geolocalización lo que impide que se localice y se ubique de forma espacial el área más próxima al sitio en el que se encuentra la emergencia.

La ausencia de un número propio puede aumentar las posibles fallas que se puedan generar al pasar la información de un operario a otro, volviendo menos eficaz y efectivo el proceso de coordinación y posterior atención de la emergencia; pues la entidad no tendría la cercanía con el ciudadano desconociendo de primera mano las necesidades del mismo.

- Manejo de las bases de datos y formularios de forma física.

El uso de bases de datos de forma física representa una desventaja para la entidad pues en el proceso de llenado de las mismas se pueden presentar diversos errores, debidos a fallas humanas comunes. La implementación de formularios virtuales le ayudará a lograr a la entidad una reducción costos, agilización de procesos, mejoramiento en la eficiencia de recolección de la información y posteriormente facilitará un análisis de la misma.

El proceso de recopilar y analizar datos para el IDIGER es una de las tareas principales y una de las más importantes durante la realización de los procesos de atención, ya que el posterior

análisis y procesamiento de la información recolectada en tiempo real permite apoyar el proceso de toma de decisiones y por consiguiente contribuir en la mejora dentro de las respuestas institucionales. El IDIGER actualmente cuenta con formularios físicos que son llevados al lugar de la emergencia y son llenados in situ, lo que posteriormente requiere que la información recolectada sea digitalizada y analizada lo que genera un doble proceso, y consecuentemente demoras en la gestión del proceso de atención.

- Falta de estandarización de las bases de datos.

Tras una búsqueda de información en bases de datos virtuales se encontró que tanto el IDIGER como demás entidades adscritas como servicio de emergencia no cuentan con bases de datos estandarizadas, y la mayoría de ellas cuenta con registros vacíos de información relevante como: el componente espacial de la emergencia, lo que produce inconvenientes en el proceso de la información, así como en el uso que se le pueda dar posteriormente a la misma dentro de los procesos de conocimiento y prevención del riesgo.

- Ausencia de variable geográfica en las bases de datos.

La ausencia del factor geográfico en los registros de información de los formularios y base de datos del IDIGER es uno de los factores más relevantes a tener en cuenta y a mejorar, ya que si se conoce el territorio, la forma en cómo se están comportando los diferentes fenómenos sobre el mismo y como la población se relaciona con dicho territorio y como puede verse afectada por los fenómenos que se presenten allí, es de gran importancia al momento de atender emergencias; también cobra gran relevancia en los diferentes procesos de la Gestión del Riesgo de Desastres (GRD) como: en el conocimiento y prevención del riesgo.

- Desaprovechamiento de la información generada por usuarios de internet en tiempo real.

Como se ha mencionado en varias ocasiones en la actualidad vivimos en un mundo globalizado interconectado en donde a diario se están generando grandes volúmenes de todo tipo que va desde como una simple fotografía capturada con la cámara de un teléfono celular, puede contener datos de posicionamiento proveniente de satélites, datos climáticos, mensajes de texto, Tweets, email y cientos de llamadas, todo en tiempo real; además y gracias a los diferentes avances tecnológicos dentro del campo de las ciencias de la información, estos han permitido reducir los costos de almacenamiento de datos masivos, impulsando de manera gradual y exponencial la acumulación de todo tipo de datos valiosos para las empresas y las organizaciones, que son empleados para predecir los comportamientos de las personas ante diversas situaciones, así como para identificar tendencias y patrones.

La actual abundancia de datos provenientes de instrumentos de medición como: Sensores, cámaras, GPS y demás; los datos generados por personas a través de los instrumentos anteriormente nombrados se posicionarán como una de las mayores tendencias y una de las partes fundamentales del conjunto de datos disponibles a nivel global que ayudarán en la identificación y resolución de problemas, mejorando de manera eficientes los diferentes procesos y las entidades u organizaciones.

### Fortalezas

- El nivel de recepción y conexión que tiene el IDIGER frente a otras entidades y Capacidad de coordinación de servicios de respuesta.

El IDIGER al ser el encargado de coordinar las entidades encargadas que pueden o que prestan los servicios de respuesta, tiene una máxima aceptación, prioridad y apoyo frente a cualquier otra entidad en una situación de emergencia para ser capaz de prestar y coordinar los 16 diferentes servicios de respuesta básicos; en la medida en que los daños y crisis social lo demanden, dichos servicios son: salud, búsqueda y rescate, extinción de incendios, manejo de materiales y/o residuos peligrosos, evacuación, ayuda humanitaria, alojamientos temporales, agua potable, energía y gas, telecomunicaciones para la comunidad, restablecimiento de contactos familiares, saneamiento básico, manejo de escombros y

residuos vegetales, manejo de catástrofes, seguridad y convivencia, información pública y accesibilidad y transporte.

Así mismo existen entidades adscritas que ejecutan los servicios de respuesta a emergencias anteriormente nombrados a las cuales el IDIGER tiene la capacidad y deber de coordinar al momento de atender una emergencia; existen las entidades distritales, regionales y nacionales y las entidades privadas.

Para atender específicamente inundaciones, el IDIGER se encarga de definir cuáles de los servicios de respuesta son activados para hacerle frente a la emergencia, comúnmente se activan: seguridad y transporte, salud, evacuación asistida, ayuda humanitaria, alojamientos temporales, agua potable, telecomunicaciones para la comunidad, saneamiento básico, manejo de escombros y obras de emergencias y seguridad y convivencia.

Para este tipo de emergencia los ejecutores de los servicios de respuesta también se dividen en: distritales, nacionales o regionales y privados; dentro de los distritales se encuentran: secretaria Distrital de Movilidad, secretaria Distrital de Salud, IDIGER, Secretaria Distrital de integración social, empresa de acueducto de Bogotá, Unidad administrativa de especial de servicios públicos, secretaria distrital de seguridad convivencia y justicia, entre otros; dentro de los privados: Codensa y Operadores de red de banda ancha, datos, telefonía fija y móvil; Dentro de los nacionales y regionales no se encuentran responsables directos de la ejecución de servicios de respuesta, ya que solo se recurren a instancias nacionales directas en caso de calamidad pública.

- Capacidad para coordinar acciones o ejecutarlas frente a emergencias por inundaciones.

Como se nombró anteriormente el IDIGER tienen dos funciones específicas: la de coordinar y también la de ejecutar, es decir que la entidad está en la capacidad de coordinar a los servicios de respuesta a emergencia, así como coordinar las diferentes entidades ejecutoras de respuesta para que atiendan la emergencia sin comprometer recursos e insumos propio,

pero también, el IDIGER puede actuar como ejecutor de servicios de respuesta por lo tanto tiene funciones como servicio de respuesta y puede atender las diferentes emergencias con recursos e insumos propios.

- Integración de otras entidades de carácter mixto: Servicios público-privados.

El IDIGER cuenta con un Plus importante, ya que está en la capacidad de coordinar y conectar con entidades de carácter mixto es decir: publico privadas que a su vez pueden o no ejercer atención a desastres como servicio de respuesta, sin embargo, deben acatar las funciones de dichos servicios de emergencia con el objetivo de optimizar dichos servicios de respuesta, entidades tales como: Defensa Civil, Bomberos Voluntarios, Cruz Roja, y empresas prestadoras de servicios públicos básicos, como: acueducto y alcantarillado, energía y alumbrado público, gas natural y demás, y empresas prestadoras de servicios de telefonía e internet.

Las funciones de respuesta ejecutadas de manera óptima son: planeación y manejo general de la emergencia, evaluación de daños, riesgos asociados y análisis de necesidades, logística, telecomunicaciones para la respuesta, aspectos financieros y jurídicos.

- a. Es función de respuesta de planeación y manejo general de la respuesta la de formular el plan de acción de la emergencia, correspondiente al tipo de evento y los daños generados por el mismo tanto a bienes físicos como daños a personas, así como, coordinar la ejecución eficaz, eficiente y oportuna de los servicios de respuesta, y el seguimiento de la emergencia hasta su respectivo cierre.
- b. Es función de la evaluación de daños, riesgos asociados y análisis de necesidades el de generar, analizar y recopilar la información sobre la magnitud de los daños, la afectación del evento a la sociedad, así como la cuantificación de las necesidades inmediatas y futuras con el fin de activar o suspender servicios de respuesta. Cabe destacar que la evaluación de los daños puede ser requerida en distintos momentos y en distintos detalles.

- c. Es función de la información pública informar de manera oficial, publica, veraz y masiva sobre las causas, efectos y acciones realizadas durante el evento, así como las recomendaciones.
- d. Es función de la logística asegurar el aprovisionamiento y distribución de suministros, servicios, montaje y desmontaje de instalaciones y equipos necesarios para atender la emergencia, así como la administración de dichos recursos y servicios.
- e. Es función del servicio de telecomunicaciones mantener de manera efectiva las comunicaciones remotas entre y para los diferentes ejecutores de los servicios de respuesta, instalando y asegurando la dotación necesaria de equipos de comunicaciones para los vehículos, el personal y las instalaciones entendidas como puestos de mandos unificados.
- f. Es función de los aspectos financieros la de analizar las necesidades financieras, la identificación de fuentes potenciales de financiamiento y la gestión de los recursos, también se encarga del seguimiento del destino de inversión de los recursos y garantiza la disponibilidad del mismo para la prestación y correcto funcionamiento de los diferentes servicios de emergencia.
- g. Son funciones de los aspectos jurídicos velar por el cumplimiento del marco jurídico en la prestación los diferentes servicios de respuesta, así como la formulación de mecanismos de soporte jurídicos que apoyen las diferentes ejecuciones, en su gran mayoría contempla acciones burocráticas y administrativas de orden legal.
- Personal capacitado para realizar labor en campo; en el lugar de la emergencia y Escalabilidad de la emergencia.

La subdirección de manejo de emergencias y desastres cuenta con un equipo que se encuentra totalmente capacitado para hacer el reconocimiento y así mismo escalar la emergencia frente a inundaciones.

### Oportunidades

- Generación de cartografía.

Dentro de las oportunidades que se evidenciaron en pro a mejorar la subdirección de manejo de emergencias y desastres del IDIGER, se determinó que la generación de cartografía propia es una de las ventajas más importantes que tendría la entidad, teniendo en cuenta que la cartografía es la representación bidimensional del territorio, que está orientada hacia la representación temática de fenómenos que ocurren en el mismo, ya que el IDIGER cuenta con la información espacial e insumos suficientes para generar cartografía completamente actualizada de los fenómenos naturales, específicamente inundaciones, lo que le permitiría tener un completo conocimiento de las zonas afectadas, así como un registro histórico de afectaciones por zonas, lo que le permitiría llevar a cabo un análisis espacial identificando las posibles amenazas y por consiguiente generando soluciones adecuadas de los mismos, teniendo en cuenta que la información geográfica es uno de los objetos de mayor demanda no solo para las empresas y ciudadanos, sino que también para entidades como el IDIGER.

- Georreferenciación de las casas y zonas afectadas por inundaciones.

El proceso de georreferenciación es una de las mejores oportunidades que tiene el IDIGER para complementar la creación de cartografía, ya que se puede, mediante este proceso, determinar de forma clara y asertiva el número de unidades habitacionales en riesgo o afectadas por un evento de inundación, y tener un registro completo de las personas que necesitan ayuda, y así mismo generar un registro histórico de afectaciones, y plasmarlo todo en conjunto en mapas temáticos virtuales que estén disponibles al público, y que sirvan como base cartográfica de diferentes estudios, trabajos de investigación y esquemas o planes de ordenamiento territorial distritales.

Actualmente el IDIGER no cuenta con un sistema de georreferenciación dentro de los bases de datos ni formatos, lo que impide de manera significativa poner la información espacial en cartografía temática, que es de gran utilidad para la entidad, la georreferenciación es una de las herramientas más importantes que se basa y va a la vanguardia de los diversos desarrollos tecnológicos de los sistemas de posicionamiento global, mejor conocidos como GPS, y también de los sistemas de información geográfica, que permiten a las entidades contar con nuevos y modernos instrumentos que dan a conocer la distribución espacial de los fenómenos naturales que afectan a la población que reside en determinado territorio.

- Fortalecimiento y/o creación de sistemas de monitoreo electrónicos y automatizados.

La subdirección de manejo de emergencias y desastres puede recurrir a los nuevos sistemas y dispositivos de detección de inundaciones desarrollados en tiempo recientes que puedan generar información relevante en tiempo real, y alimentar las bases de datos y generar insumos con el objetivo de mantener actualizada la cartografía. También puede valerse de los sistemas que están actualmente implementados como las estaciones meteorológicas del IDEAM, para medir y monitorear la precipitación, lo anterior para generar alertas tempranas, siendo este uno de los elementos principales que contribuyen de manera efectiva a evitar la pérdida y afectación de bienes materiales y vidas humanas, de las poblaciones vulnerables o afectadas por inundaciones.

Los sistemas de alertas tempranas pueden ser aplicados a cualquier tipo de amenaza que pueda afectar a determinada población, ya sean fenómenos naturales o antrópicos; se realiza mediante un monitoreo constante de variables o indicadores que son específicas a cada fenómeno, para este caso: inundaciones, para la entidad se propone implementar una de las dos modalidades de alertas tempranas: estrategias automáticas de recolección, ya que se reducirían los tiempos y se automatizarían varios procesos en menos pasos, con el objetivo de pronosticar y hacer seguimiento de manera eficaz de los diferentes fenómenos naturales y fortalecer la capacidad del IDIGER para hacer frente a la atención de la emergencia.

Frente a inundaciones los sistemas de alertas tempranas comprenden un grupo de herramientas que facilitan a la población medidas de reacción frente a posibles eventos de inundación, que pueden estar comprendidos en la construcción de una red de estaciones hidrometeorológicas, o la implementación y actualización de detectores de inundaciones, que proporcionen información en tiempo real, y que mediante la implementación de diferentes Softwares y modelos predictivos de información, disminuyan los riesgos provocados por las inundaciones en Bogotá. (Losada, Rubiano, & Muñoz, 2016).

- Articulación de los consejos locales.

En Bogotá cada alcaldía local cuenta con un sistema local de gestión de riesgos y cambio climático así mismo cada una de esas dependencias locales cuenta con un plan de acción, estos consejos locales se encargan de informar a la comunidad de sus respectivos sectores sobre los riesgos, es decir que se encargan de los aspectos relacionados con conocimiento y prevención del riesgo, para el IDIGER estos consejos se presentan como una oportunidad ya que se pueden articular de manera efectiva bajo unos mismos estándares en pro de mejorar la recolección y distribución de la información desde y hacia las comunidades.

#### Amenazas:

- Las entidades que conforman y ejecutan los servicios de respuesta a emergencias son independientes.

Cada una de las entidades que conforman los servicios de respuesta a emergencias tanto públicas como privadas, son independientes una de la otra y no existe una jerarquía entre ellas y el IDIGER no tiene ningún tipo de control sobre las decisiones y gestión de las mismas, sin embargo y como se explicó en los incisos anteriores el IDIGER si tiene prioridad entre las entidades, sin embargo representa un riesgo debido a que ante una situación de emergencia como una inundación, se debe coordinar cada servicio de respuesta por separado y se debe tener en cuenta dicha independencia dentro de los procesos de atención, pues la

gestión de cada servicio de respuesta ante un desastre se puede llevar a cabo de forma diferente, sin embargo deben seguir los procesos y directrices del IDIGER.

- Recursos

Los recursos y de donde salen para hacer frente a una situación de emergencia es una de las amenazas para el IDIGER ya que no es un factor que dependa de la entidad como tal, sino que depende del presupuesto distrital, y de la asignación del monto por parte del mismo.

- Capacitación inadecuada de personal externo frente a la atención y manejo de la emergencia: inundaciones.

Al momento de atender una emergencia, no todo el personal pertenece al IDIGER, también existe capital humano que pertenece a los demás servicios de respuesta, por consiguiente, cada vez que se presenta una situación de emergencia, se debe capacitar al personal con el fin de que se cumplan los estándares propuestos para llenar los formatos, o manejar los recursos virtuales con el fin de estandarizar la información y gestionar un mejor servicio de respuesta a emergencias, lo que quita tiempo y retrasa los procesos de atención, así mismo, se pueden generar errores y vacíos en los datos, entorpeciendo el proceso de ayudas humanitarias, atención prioritaria, y censo de daños de la población afectada.

- Participación de las entidades de acuerdo a las particularidades de la emergencia.

No todas las entidades están preparadas, o se requieren para manejar una emergencia por inundación, por lo cual la coordinación es clave, pues no se deben tomar a la ligera los recursos de los servicios de respuesta cuando no son requeridos, debido a que cada entidad aporta recursos propios para hacerle frente a las emergencias por inundaciones.

- Participación social y comunitaria en la gestión de riesgos y cambio climático, mediante medios electrónicos.

La participación de la comunidad es uno de los factores clave dentro de la participación de la Gestión del Riesgo de Desastres, hoy en día casi todas las personas cuentan con un dispositivo móvil con acceso completo a internet, ya que existe una consolidación de la tecnología, la diversidad de satélites artificiales, la amplia y variada gama de sensores creados para la captura de datos geográficos como: datos climáticos, imágenes de la superficie terrestre, temperatura, velocidad de los vientos, datos de precipitación entre otros. sin embargo, otra de las grandes cantidades de datos geográficos que se generan de manera cotidiana se da mediante la interacción de las personas con las redes sociales, sitios web, juegos, ente otros, ya que muchas de dichas interacciones permiten compartir la ubicación geográfica de forma inmediata, permitiendo ser mostrada en mapas en tiempo real y de forma permanente, todos los datos anteriormente nombrados generados por la huella digital de las personas en los medios virtuales, permite que el manejo de datos masivos Big Data, sirva para predecir fenómenos y obtener la ubicación exacta de los dichos fenómenos. En este caso se considera una amenaza ya que la entidad como tal no controla todas y cada una de las interacciones de las personas con los medios virtuales, por lo que se considera pertinente una debida educación de la comunidad.

### 5.2.1 ESTRATEGIAS F.O

Dentro de las estrategias que se consideran dentro de la transversalidad fortalezas oportunidades, se encontró que la obtención de información e insumos de diferentes entidades para alimentar la cartografía es fundamental, ya que al tener fácil apoyo de las demás entidades, también se puede lograr una fácil participación y colaboración activa por parte de las mismas, que permita generar una completa estandarización de los datos, reduciendo así de manera significativa los tiempos de procesamiento y análisis de dicha información.

También, la generación de rutas de atención efectivas mediante el aprovechamiento de la georreferenciación, es uno de las mayores oportunidades a tener en cuenta apoyado en la ayuda de medios digitales y sistemas de información geográfica, ya que en la actualidad un SIG permite combinar e integrar gran cantidad de información variada que no es fácilmente

interpretable por otros medios, ya que este combina información proveniente de cartografía base, con otra u otras capas de información de cualquier tipo, ya sea estadística, matemática, temática entre otras, haciendo de esta manera que las consultas sobre el territorio se puedan realizar de manera rápida y que puedan aportar información necesaria, versátil y mucho más rica para el usuario u operador. En este caso, para determinar las rutas de atención más adecuadas al momento de hacerle frente a emergencias por inundaciones.

Otra de las oportunidades que se pueden evidenciar es la reducción de los tiempos de capacitación a personal frente a una emergencia, ya que como se nombró anteriormente no existe una jerarquía entre entidades, y cada entidad dispone de recursos tanto económicos como humanos propios por lo que al momento de atender una emergencia se debe capacitar a todo el personal disponible para lograr una atención integral, lo que se puede lograr mediante la continua capacitación del personal y la reducción de rotación del mismo por parte de las entidades, sin embargo el IDIGER como entidad encargada de coordinar, puede emitir tipos de capacitación, las cuales serán un estándar para todas las entidades; y mejorar los tiempos de atención enfocándose en la Capacitación de personal enfocado al área de las TIC's.

### 5.2.2 ESTRATEGIAS D.O

Dentro de las estrategias que se consideraron entre la transversalidad debilidades oportunidades, se determinó de la creación de una línea de atención directa, es una de las formas más efectivas con las que contaría el IDIGER, para ser una entidad más cercana al ciudadano conociendo de primera mano las necesidades del mismo, obteniendo directamente de la ciudadanía información primaria necesaria para la creación y mantenimiento de mapas de calor de focos de inundaciones en Bogotá, presentando así una oportunidad de generación de cartografía propia y actualizada; la cual estaría a disposición de la entidad en cualquier momento, para llevar a cabo análisis espaciales de cualquier nivel y de cualquier fenómeno, razón por la cual se deben manejar y elaborar de bases de datos electrónicas que permitan un fácil procesamiento y análisis de toda la información que alimentaría la cartografía y a la entidad.

Así mismo, es necesario lograr una adecuada estandarización de las bases de datos, ya que la información debe estar disponible no solo para el IDIGER sino para cualquier entidad y viceversa con el objetivo de mejorar los tiempos y la gestión de una emergencia por inundación, no obstante se debe tener en cuenta que se deben integrar los componentes X e Y a las bases de datos, pues como se ha venido hablando el componente geográfico es lo que le da el plus a las empresas o entidades, como es el caso de Google, Amazon, Facebook entre otros, que han explotado la relación entre el componente geográfico y el almacenamiento de estos mismos datos con el fin de determinar patrones que facilitan la aplicación e interpretación de algoritmos en correlaciones de datos para predecir necesidades, y luego aplicar dicho conocimiento de necesidades para generar productos y servicios.

En el caso del IDIGER la aplicación del conocimiento y análisis de la información se empleará para el mejoramiento de los tiempos de atención, otro factor a considerar es la adición de un componente Z en el conjunto de datos geográficos, pues en grandes ciudades donde la propiedad horizontal es predominante se requiere saber, desde que piso se está generando la información y así también aprovechar la información obtenida de los datos generados en las redes sociales, partiendo las metodologías de aplicación del Big Data para la atención a inundaciones en otras ciudades: Chennai.

### 5.2.3 ESTRATEGIAS F.A

Dentro de las estrategias transversales entre fortalezas y amenazas se encontró que Incrementar y mejorar la capacidad de organización y coordinación entre las entidades y el IDIGER, es uno de los factores a mejorar, pues las demás entidades son externas al IDIGER, no existe una jerarquía por lo tanto se pueden generar retrasos en los tiempos de atención, mejorar la coordinación y cómo se gestiona es fundamental, también en el lugar de la emergencia es importante utilizar formatos digitales formatos digitales amigables con el usuario de fácil llenado, ya que se reducen costos y se mejora la calidad de la información, y así mismo se reducen los tiempos de análisis e interpretación de la misma.

La capacitación de la comunidad para incrementar la participación activa de la misma en redes sociales, mediante propagandas y actividades educativas es uno de los factores más importantes ya que la comunidad es el foco principal generador de información vital para los procesos del IDIGER, así mismo se debe procurar que los recursos se agoten en primera instancia de créditos contingentes y pólizas de seguro, pues los recursos propios deben mantenerse para cubrir externalidades de la emergencia, como lo son los salarios, mantenimiento de equipos e instrumentaria.

#### 5.2.4 ESTRATEGIAS D.A

Dentro de las estrategias transversales debilidades amenazas se encontró que la reducción de los tiempos de atención por entidades se debe disminuir, mejorando la gestión, la colaboración y la capacitación del personal frente al llenado de bases y manejo de medios electrónicos y virtuales, así mismo mantener informada y capacitada a la comunidad frente al manejo y utilización de los nuevos recursos virtuales como la creación de un “Hashtag” o demás comportamientos comprendidos dentro de la huella digital de dichas comunidades con el fin de aprovechar la información generada en las redes sociales.

### 5.3 Guía metodológica para la priorización de zonas afectadas por inundaciones

La siguiente guía presenta los procedimientos necesarios para la priorización de zonas afectadas por inundaciones a través del uso de los datos de las redes sociales, específicamente de Twitter. Esta se divide en dos pasos metodológicos importantes que son: (1) la recolección, filtrado y clasificación de datos en Twitter y (2) el mapeo de los datos Big Data provenientes de las redes sociales. En cada uno de ellos se proponen las herramientas y técnicas más apropiadas para la priorización. Cabe aclarar que para el paso (2) los procesos se aplicaron al barrio el Tintal, ubicado en la ciudad de Bogotá.

## 1. Recolección, filtrado y clasificación de datos en Twitter.

Para recolectar los Tweets relacionados con las inundaciones ocurridas en Bogotá o en un lugar determinado, se debe considerar los hashtags o las cuentas oficiales que dan la información de lo ocurrido. Los hashtags pueden ser palabras claves que generen popularidad entre los usuarios con el fin de que la información producida se enfoque en un tema específico como lo son las inundaciones.

En primer lugar, Twitter ofrece una herramienta denominada API (Application Program Interface) con la cual se pueden recolectar Tweets en tiempo real (streaming). Esta API, es gratuita, y sólo se requiere abrir una cuenta en el portal del desarrollador y obtener un usuario y contraseña que habilita al desarrollador a usar dicha herramienta.

Para hacer uso de esta API, hay varios lenguajes de programación que pueden consumir los servicios de streaming. Entre los más populares están Python y Java. Un ejemplo de cómo consumir los servicios de la API de Twitter es mediante el uso del lenguaje de programación Java. Dicho lenguaje se puede programar en una plataforma de desarrollo de código abierto, denominada Eclipse. Mediante Eclipse se tiene un marco de trabajo y un compuesto de servicios que permiten la construcción del entorno de desarrollo de los datos de entrada. El anexo1 muestra el entorno de trabajo de Eclipse:

Este entorno consta de 3 partes principales, la ventana más grande se le conoce como “Editor” y es donde desarrolla o codifica las funcionalidades requeridas, la ventana vertical del lado izquierdo permite navegar por el árbol de directorios de los proyectos y la ventana horizontal de la parte inferior muestra las tareas pendientes insertadas por el usuario; cabe aclarar que estas ventanas se llaman “Vistas”.

Existen dos tipos de barras de herramientas en el espacio de trabajo de Eclipse que son: la barra de herramientas principal y la barra de Perspectivas. En la primera hay accesos directos a operaciones básicas como abrir, guardar, editar entre otras y en la segunda se encuentran las perspectivas que se usan en el proyecto.

En primer lugar, se debe tener en cuenta las dos secciones más relevantes de una clase Java: la primera que define las variables globales del programa y la segunda que define los parámetros de conexión a la API de Twitter (ConsumerSecret, AccessToken, AccessTokenSecret, IncludeEntitiesEnabled) (ver anexo 2). Lo anterior, con el fin de recuperar Tweets en tiempo real, relacionados con cuentas de entidades vinculadas a la atención de emergencias, de medios de comunicación y hashtags de uso común en emergencias. Estos Tweets se guardan en archivos planos (ej: fomartos .json, .csv) y en base de datos denominada Mongo DB.

La configuración de los hashtags, cuentas de entidades y medios de comunicación que se recuperan por streaming en tiempo real, se hacen teniendo en cuenta unas coordenadas geográficas específicas. En cuanto a los Hashtags se propone la generación y distribución en Twitter de más hashtags como #InundacionBogota #InundacionBOG #Inundacionen\_(Nombre de un barrio)\_ . Al correr el proceso, se obtiene un archivo plano de los tweets recolectados en formato .json y se guardan los Tweets en la base de datos Mongo DB para verlos de forma más organizada.

Los Tweets recolectados se pueden analizar (filtrar y clasificar) con la herramienta Weka, la cual es una colección de algoritmos de aprendizaje automático para actividades de minería de datos, tiene métodos para la preparación, clasificación, regresión, agrupación de datos y demás, es de uso público y se puede descargar en: <https://www.cs.waikato.ac.nz/~ml/weka/>.

La figura 14 muestra el selector GUI de Weka que proporciona un punto de partida para lanzar las principales aplicaciones GUI y las herramientas de soporte de Weka. Este está compuesto por 4 opciones que son: el explorador, el experimentador, KnowledgeFlow, Workbench y SimpleCLI.

Figura 14. Selector GUI de Weka



Para este caso se propone usar la interfaz de Explorer ya que permite hacer operaciones sobre un conjunto de datos como el pre procesamiento de datos, aplicación de filtros, clasificación, clustering y otros. Es necesario considerar que los datos que fueron recolectados y guardados en un formato .json, deben convertirse, mediante un programa existente en cualquier lenguaje como Python, a uno los formatos que se trabajan en Weka como:

- ARFF: (Attribute-Relation File Format) es un archivo de texto ASCII que describe una lista de instancias que comparten un conjunto de atributos. Los archivos ARFF tienen dos secciones distintas, la primera sección es la información del encabezado, que contiene el nombre de la relación, una lista de los atributos (las columnas en los datos) y sus tipos; y la segunda es la información de datos.

Al seleccionar la opción del Explorador automáticamente se abre una pestaña de pre procesamiento (ver anexo 3), que permite elegir y modificar los datos. Cuando se carga un archivo .arff se muestra información sobre los atributos e instancias de los datos así como un histograma en la esquina inferior derecha y se puede empezar a transformar los datos mediante la aplicación de filtros.

En el encabezado del archivo .arff se especifican los atributos y sus tipos por ejemplo numérico, texto, fecha y otros; en el anexo 4 hay un ejemplo aplicado a datos climáticos y se puede ver que los atributos se denotan con @attribute. Para efectos prácticos en este proyecto

se pueden utilizar atributos como las coordenadas de los Tweets, el ID, nombre del usuario de Twitter, el texto que contiene información que el usuario envía frente a un tema, la fecha en que se creó el tweet, el número de Retweets y otros que ayuden a dar la información necesaria durante una situación de emergencia.

Teniendo los datos cargados se debe realizar una preparación de los mismos aplicando la opción de `StringToWordVector` para ir quitando las palabras vacías *Stopwords* como artículos, pronombres, preposiciones, etc. e ir agrupando las palabras que realmente importan.

La pestaña de clasificación, sirve para entrenar y probar esquemas de aprendizaje que clasifiquen o realicen regresión. El algoritmo de aprendizaje automático que se recomienda usar para hacer una clasificación de los Tweets es el algoritmo de Bosques Aleatorios (Random Forests). El resultado luego de aplicar el clasificador debe ser entrenado varias veces con las opciones de prueba que se establezcan hasta obtener un buen porcentaje en los aciertos de clasificación.

Para este tema de estudio los tweets se podrían categorizar teniendo en cuenta el significado de las palabras, por ejemplo si se consideran palabras importantes como inundación, ayuda, auxilio, emergencia, donaciones, afectados, puede formarse una categoría de mayor prioridad que contenga una etiqueta como Necesidad de Ayuda o Activación de Protocolo de Emergencias. Cuando existen palabras que informan sobre la emergencia se puede denotar con la etiqueta de Fuente de Información o Referencias Informativas, las palabras que no tienen relación con la situación pueden etiquetarse como Información Errónea o Ruido, las opiniones y comentarios pueden tener una etiqueta de Otros u Opiniones Públicas.

Es preciso señalar que para realizar todas las operaciones descritas anteriormente se necesita de un científico de datos o personas especializadas en el tema que tengan la capacidad de analizar e interpretar grandes bases de datos en poco tiempo.

## 2. Mapear los datos Big Data provenientes de las redes sociales.

Como ya se nombró anteriormente para la subdirección de manejo de emergencias y desastres del IDIGER es de vital importancia dentro de la transversalidad de las estrategias propuestas, representar la información obtenida de los datos generados por las redes sociales de forma gráfica (mapas), mediante la utilización de herramientas cartográficas como los SIG que permitan generar un análisis detallado de la emergencia, incorporando el elemento espacial, mediante la descripción del territorio y por lo tanto las dinámicas que se estén generando sobre el mismo y la manera en como la población interactúa y es afectada durante dicho fenómeno entendido como una emergencia por inundación.

Para lo cual fue necesario realizar una aproximación y una revisión de la metodología empleada en el uso y análisis de Twitter durante la Inundación de Chennai (2015), para la gestión de desastres, ya que fue la metodología que más se aproximó a las necesidades de la subdirección de manejo de emergencias y desastres del IDIGER, pues emplea de manera exitosa el uso de las redes sociales, específicamente Twitter, de la cual se obtiene información primaria de la población afectada, así como la agregación del componente espacial, es decir de las coordenadas x e y a la información, permitiendo que la representación cartográfica de los datos sea más efectiva y confiable, pues hoy en día cualquier tipo de dato es georreferenciable, y la mayoría de los datos de los cuales se desglosa la información pertinente de internet o de las redes sociales cuentan con dicho atributo ya integrado.

No obstante, la subdirección de manejo de emergencias y desastres del IDIGER no tiene incorporado el uso de la información obtenida de los datos generados en las redes sociales, ni maneja una cartográfica propia que facilite la gestión de la atención dentro de las respuestas institucionales como se vio evidenciado en el análisis de la estructura orgánica y subprocesos en la subdirección que dio como resultado una matriz de análisis DOFA, razón por la cual durante el desarrollo de este inciso se proponen métodos para priorizar zonas afectadas por inundaciones haciendo uso de herramientas cartográficas.

Las aplicaciones del Big Data son múltiples, contribuyendo a impulsar los grandes cambios tecnológicos de la era moderna como: Smart Cities y el internet de las cosas, resolver catástrofes y el cambio climático, ya que su objetivo principal es compartir y analizar datos que puedan contribuir al desarrollo de la sociedad, así mismo fomenta la innovación y el desarrollo.

La cartografía inteligente, representada como la tecnología que traslada los datos de la vida cotidiana bien sea en tiempo real o no, a los mapas, que son de fácil visualización, con el fin de mejorar la toma de decisiones a partir de un cumulo de información en su mayoría compleja; mediante la cartografía inteligente se pueden elaborar mapas conforme a las necesidades, apoyados en los sistemas de información geográfica (SIG).

Dichos SIG permiten visualizar numerosos datos atendiendo al mismo tiempo a las cuatro dimensiones del Big Data mejor conocidos como las cuatro “V”: valor, volumen, veracidad y variedad, ya que en definitiva los mapas son una simplificación de la información de manera gráfica y permiten entender la realidad con tan solo un clic; la clave es que los SIG utilizan la geolocalización, es decir , se apoya en las coordenadas geográficas, como llave primaria entre los datos que van a ser materia de análisis así como de representación y el territorio, por lo tanto la clave principal es incorporar una columna con un atributo estrechamente relacionado a la localización, es decir, elementos similares a la ubicación geográfica, como la localidad, el número de la manzana y la placa catastral.

Con base en lo anterior, las herramientas que se pueden utilizar para analizar y visualizar los datos están dadas en función de quien va a generar los datos y se puede hablar de tres tipos:

1. Privadas: hace referencia a empresas privadas como Google, que a través de sus servicios: Google Maps, que permiten trabajar online tanto con los mapas, como con los datos.
2. Públicas: todas aquellas entidades que usen y dispongan de una infraestructura de datos espaciales (IED), que puedan homogenizar, facilitar y plasmar los datos de

forma gráfica, y así mismo ponerlos al servicio de los ciudadanos a través de geoportales como el SIG OT o el IDECA.

3. Ciudadanas: se abre paso con aplicaciones como Openstreetmap una herramienta colaborativa donde los datos son de carácter público y abiertos para poder ser mejorados entre todas las personas con acceso a ellos.

Debido a los sistemas de información geográfica, que permiten trabajar mediante capas de información, es posible trabajar sobre una base topográfica robusta, sobre la cual se puede seguir incorporando datos en formato vector, así mismo se puede incorporar datos y capas geográficas mediante servidores a través de formatos WMS (Web Map Server), datos que se encuentran directamente sobre la nube, pero que requieren ser trabajados a través de un software geográfico.

Mediante la gestión adecuada de los datos y teniendo a disponibilidad la mayor cantidad de datos geográficos de Bogotá, fue posible realizar una modelación cartográfica a partir de una muestra manejable de datos del barrio el Tintal, una de las nuevas ciudadelas ubicadas al occidente de Bogotá, y que pertenece a la Localidad de Fontibón (9); dicho sector se escogió debido a que se encuentra ubicado dentro de las tres zonas de riesgo; y que mediante la utilización procesos estadísticos de la caja de herramientas del ArcGis ( ArcToolBox<Analysis Tools< Statistics< Tabule Intersection) se pudo determinar que: de los 1.13 Km<sup>2</sup> de su respectiva área, el barrio el Tintal, tiene: 252.091 m<sup>2</sup> se encuentran en zona de Amenaza Alta; 303.800 m<sup>2</sup> en zona de Amenaza baja; 543.280 m<sup>2</sup> en zona de Amenaza Media; y el 39.078 m<sup>2</sup> restantes se encuentran ubicados en zonas de ampliación del cauce y sobre el cauce en cuestión, como se muestra en la tabla 5 y gráfico 1: el área total correspondiente al barrio el Tintal, así como, cada una de las áreas en metros cuadrados que se encuentran sobre zonas de amenaza.

Tabla 5. Área de intersección en zonas de amenaza.

Descripción de la Amenaza		Área Total B. El Tintal (m <sup>2</sup> )	Área de Intersección (m <sup>2</sup> )
1	Amenaza alta	1'136.780	252.091,79
2	Amenaza media		303.800,01
3	Amenaza baja		543.280,19
4	Ampliación del cauce		20.055,03
5	Cauce oficial		19.023,76

Gráfica 1. Área de intersección en metros cuadrados

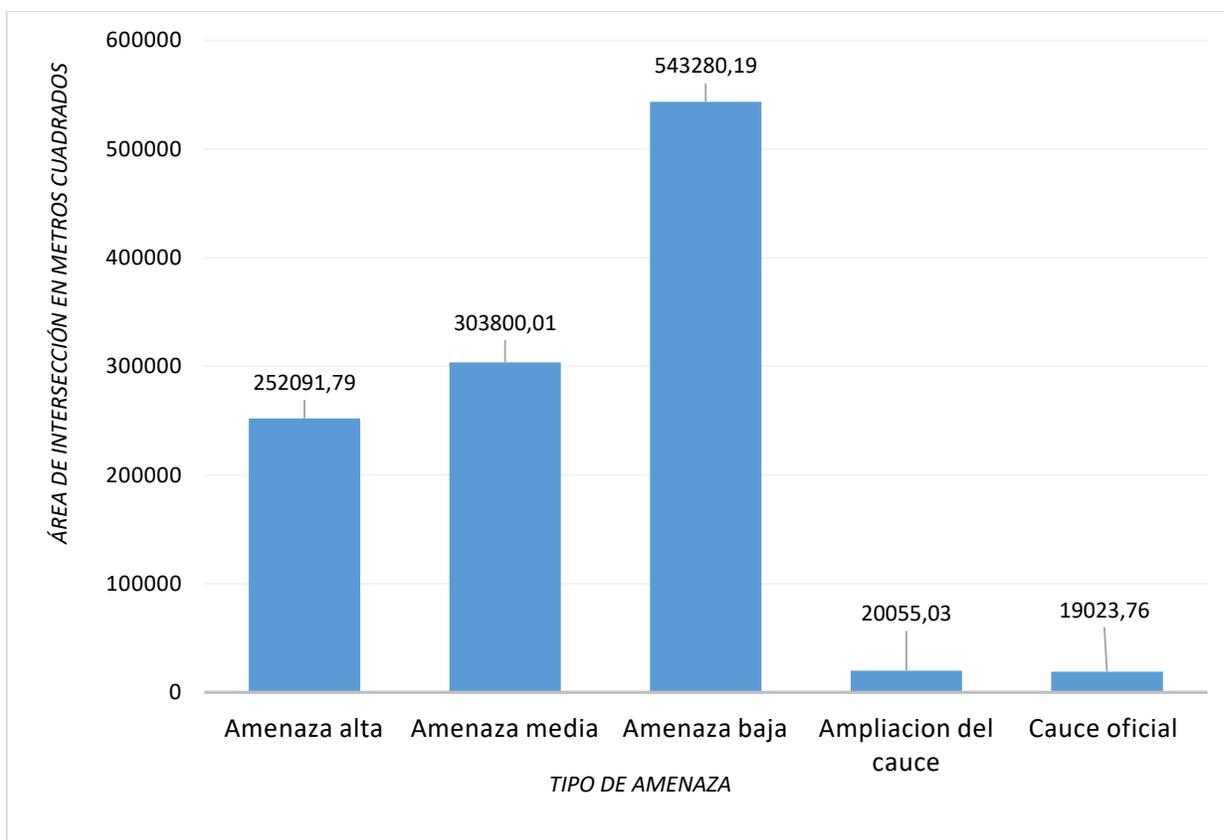
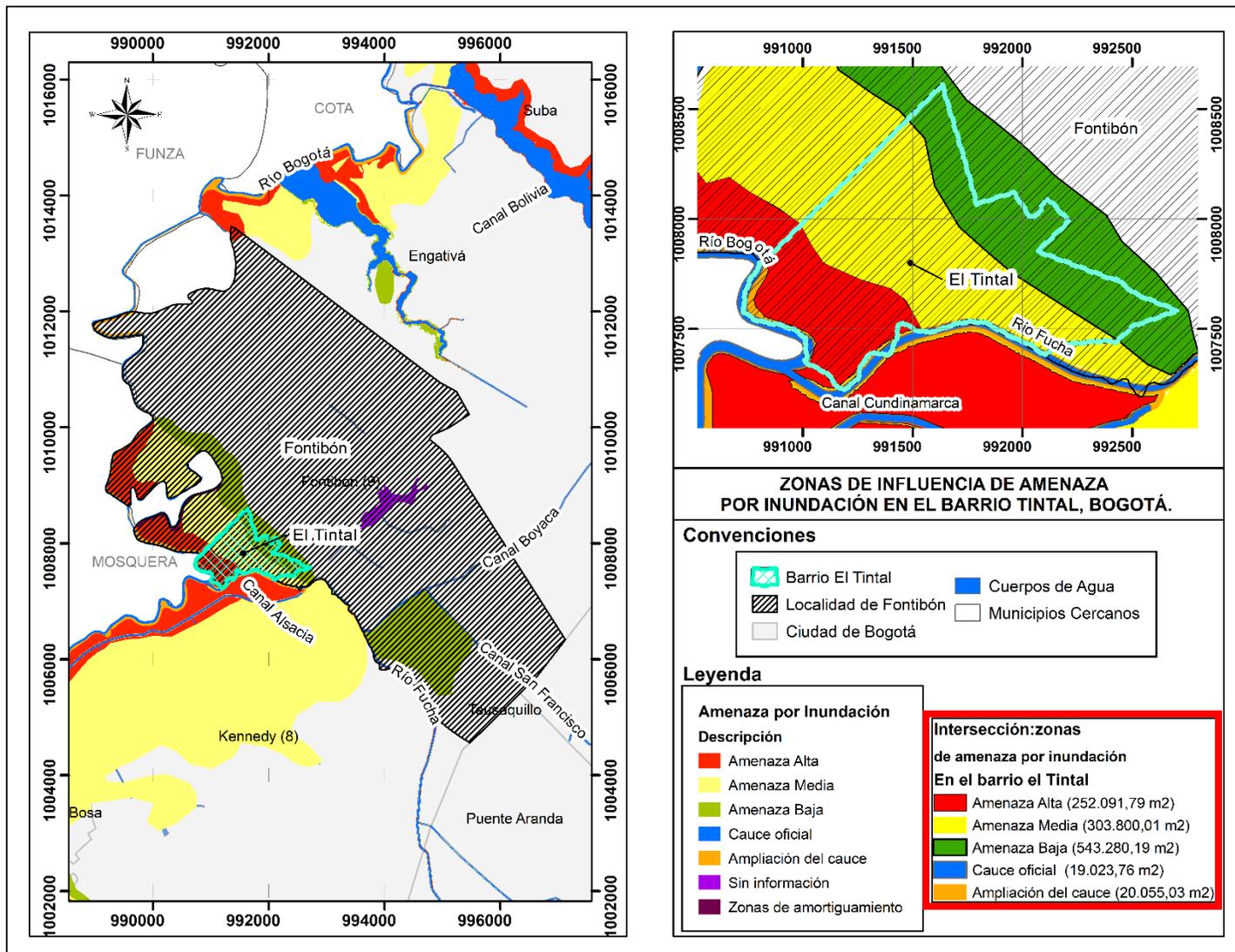


Figura 15. Zonas de intersección en áreas de amenaza



En la figura 15 se pueden observar las localidades vecinas junto con la capa de información correspondiente a la zonificación de amenaza por inundación, la información base se obtuvo de los servidores de la IDE para Bogotá, como ya se nombró anteriormente, junto con la ampliación del sector escogido para realizar la muestra de gestión de datos provenientes de redes sociales representados cartográficamente, como se mencionó antes, se escogió el barrio catastral el Tintal debido a que se encuentra ubicado sobre las tres zonas de amenaza por inundación, y su aptitud de uso de suelo tiene una configuración de modulo habitacionales.

De igual manera, la figura 16 muestra un modelo de elevación digital para la ciudad de Bogotá ya que dicho modelo es una representación visual y matemática de los valores de Z en decir los valores de altura para la ciudad de Bogotá con respecto al nivel del mar, y que permite ver y comprender de forma gráfica las zonas de amenaza por inundación (ver figura 17) pues a través de dichos modelos es posible conocer la disposición, existencia, forma y posición de los diferentes elementos que conforman el territorio geográfico y que bien pueden ser de origen natural o antrópico.

Lo anterior indica que el número de edificaciones que se encuentran en situación de riesgo por encontrarse dentro de las zonas de influencia y amenaza por inundaciones en Bogotá es de 224.585 m<sup>2</sup>; en el barrio el Tintal es de 3.267 m<sup>2</sup> lo que para efectos de este proyecto representa una muestra significativa, ya que la mayor parte de dichas edificaciones están destinadas a uso residencial, razón por la cual es de gran importancia tener una metodología de priorización en la atención a emergencias, por lo cual se propone utilizar los datos obtenidos de las redes sociales como se explicó anteriormente, agregarles un valor geográfico que hagan posible su fácil representación cartográfica, y generar un modelo de identificación de zonas afectadas “Hot Spot” : polígonos de Thiessen o interpolación IDW.

Figura 16. Modelo de elevación digital para la ciudad de Bogotá

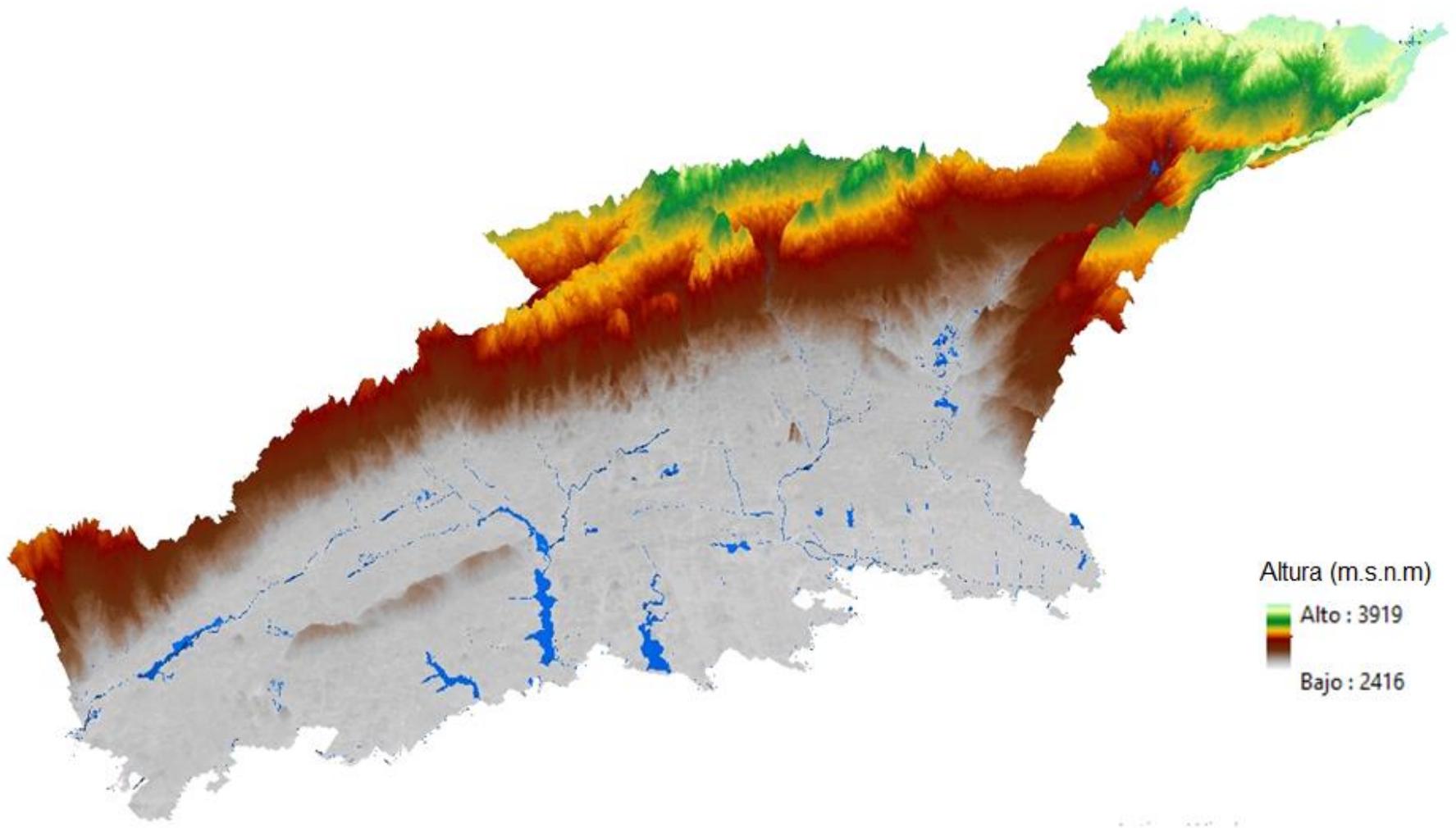
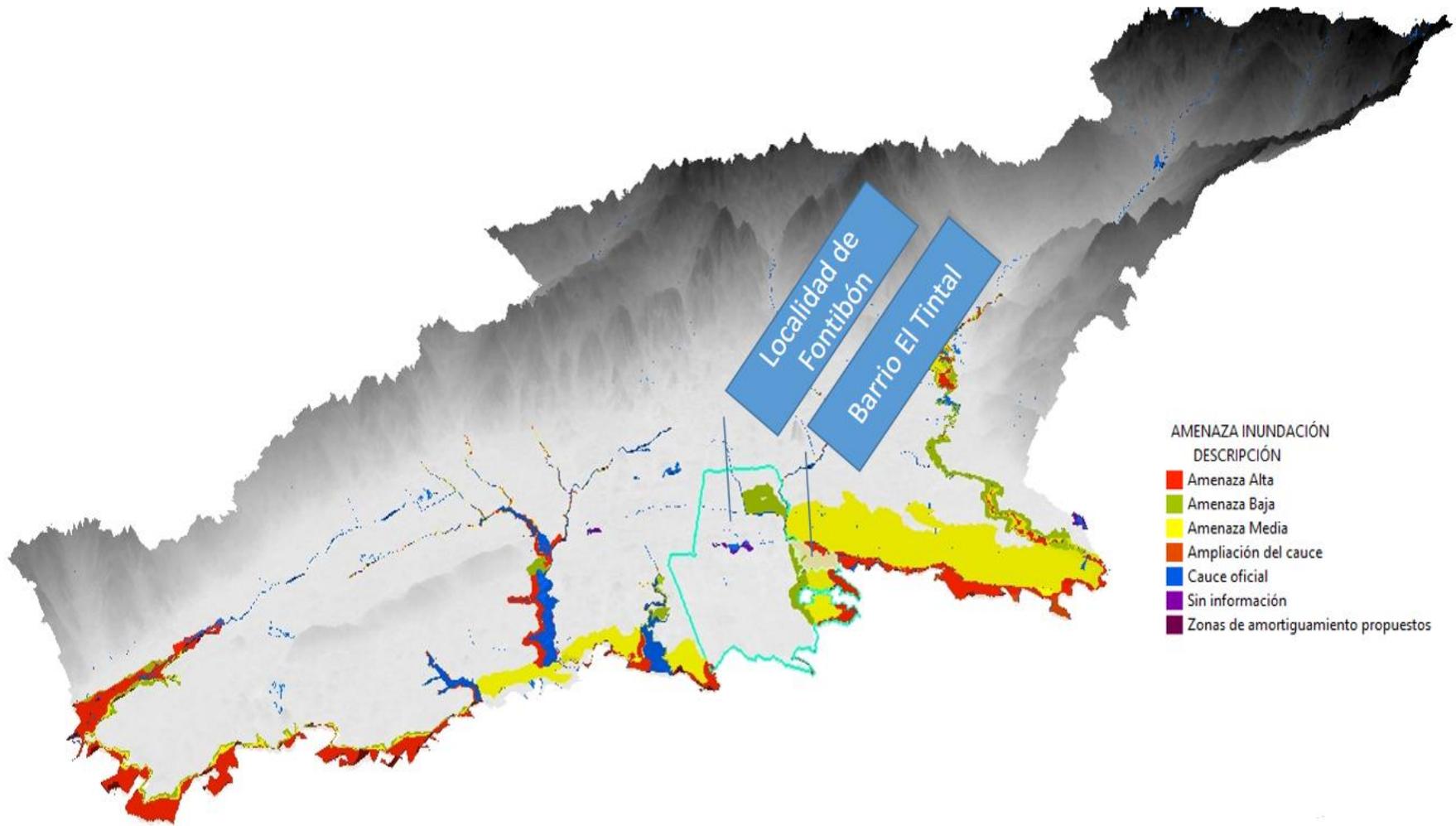


Figura 17. Modelo de elevación digital con zonas de amenaza por inundación.



Una vez realizados los procesos de organización de la información inicial en bases de datos compatibles con los sistemas de información geográfica, se puede realizar un proceso de interpolación consistente en hallar un dato dentro de un intervalo en los cuales se conocen los valores en los extremos.

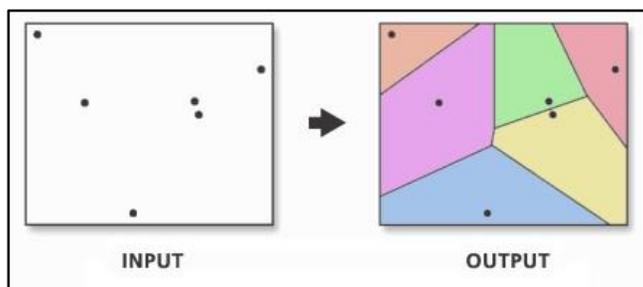
Para este caso los valores se conocen mediante la placa catastral, la dirección o el manzaneo urbano de los inmuebles afectados en los bordes extremos de las zonas de influencia por inundaciones en la ciudad de Bogotá, lo cual se puede realizar mediante la creación de polígonos de Thiessen el cual representa un método para representar áreas de influencia en unidades discretas a partir de un conjunto de puntos, así pues, el tamaño y la configuración de los diferentes polígonos resultantes dependerá de la distribución de los puntos, el método de los polígonos de Thiessen es comúnmente usado para efectuar análisis de precipitación en estudios meteorológicos, sin embargo para este caso se usaran con el adecuado shapefile de puntos a partir de coordenadas geográficas asociado a un código único georreferenciado de las viviendas en la ciudad de Bogotá como registro único que se pueden organizar en intervalos con el fin de representarlo cartográficamente.

Cada polígono de Thiessen se crea a partir de características de puntos (ver figura 18) y contiene una sola función de entrada mediante dichos puntos, así, cualquier ubicación dentro del polígono está más cerca del punto asociado; dicha metodología se usa comúnmente para dividir el área cubierta por entidades de puntos como se mencionó anteriormente de zonas próximas entre sí con valores de interpolación similares, en donde todos los puntos se triangulan en una red irregular triangulada (TIN) que cumple con los criterios de Delaunay,<sup>22</sup> “Las bisectrices perpendiculares para cada borde del triángulo se generan, formando los bordes de los polígonos de Thiessen. La ubicación en la que se intersectan las bisectrices determina las ubicaciones de los vértices del polígono de Thiessen” (ArcgisPro, 2019).

---

<sup>22</sup> “La triangulación de Delaunay es una de las triangulaciones más interesantes por ser aplicable para la resolución de multitud de problemas aparentemente sin relación entre sí, debido a sus propiedades geométricas, y por contar con algoritmos bastante eficientes para su cálculo. Todo ello además implica que hay una gran cantidad de material escrito sobre este tipo de triangulación.” (UPM, 2019).

Figura 18. Creación de polígonos de Thiessen desde entidades de puntos



Tomado de: ArcGis Pro Help: <https://pro.arcgis.com/en/pro-app/tool-reference/analysis/create-thiessen-polygons.htm>

De igual manera, otra de las herramientas mediante las cuales se puede realizar un proceso de priorización de las zonas afectadas es mediante una interpolación IDW o mejor conocido como interpolación de distancia inversa ponderada, que determina cada uno de los valores de celda mediante la combinación ponderada linealmente extraída de un conjunto de puntos, que para este caso es la muestra de viviendas que reportaron una emergencia por inundación, sin embargo para este tipo de procesos de interpolación, la variable a evaluar debe ser diferente y dependiente a la ubicación geográfica, razón por la cual y como se dijo anteriormente es necesario que la base de datos cuente con más información como el tipo o magnitud de la emergencia, otra de las razones por la cual utilizar un modelo de interpolación IDW es pertinente, es debido a que es un modelo flexible que permite establecer barreras es decir, que si existen barreras naturales como un pico de cambio en la elevación, un canal, una ampliación en el cauce y demás, dicha barrera le impedirá buscar los puntos de entrada de la muestra afectados por dicha barrera.

En la figura 19, se muestra la ampliación de la zona afectada en cuestión para efectos de la muestra, es decir la zona correspondiente al barrio el Tintal, que como se dijo anteriormente se encuentra sobre las tres zonas de amenaza por inundación, así mismo se pueden observar los Shapfiles de puntos correspondientes a las placas catastrales de las viviendas ubicadas sobre el mismo sector, y las placas catastrales correspondientes a la muestra; también se puede observar un raster correspondiente a la aplicación y utilización de la metodología de polígonos de Thiessen, el cual es un método asertivo para analizar zonas de afectación por área, ya que como se puede observar, el color rojo corresponde a zonas más pequeñas en

metros cuadrados, lo que para efectos de este trabajo se considera que son las zonas que se deben priorizar, ya que son las zonas desde las cuales se recibieron más reportes es decir que son las zonas en donde se está presentando mayor densidad de personas afectadas por metro cuadrado por el aumento del cauce normal del agua.

En el siguiente mapa (ver figura 21), al igual que en el mapa de la figura anterior, se muestra la ampliación de la zona afectada en cuestión para efectos de la muestra, es decir la zona correspondiente a el barrio el Tintal, así mismo se muestran los Shapesfiles de puntos correspondientes a las placas catastrales de las viviendas ubicadas sobre el mismo sector, y las placas catastrales correspondientes a la muestra.

También se puede observar un raster correspondiente a la aplicación de la metodología de interpolación IDW; el cual funciona mediante la distancia inversa ponderada y determina los valores de celda a través de una combinación ponderada linealmente de un conjunto de puntos de la muestra, el cual es un método asertivo para analizar zonas de afectación por tipo de emergencia, por lo que para esta muestra se utilizaron tres tipos de emergencias asociadas a las inundaciones y se les asignó un numero o código: Inundaciones (2), Encharcamientos (3) , devolución de Aguas residuales (4), como se puede observar en las clases del raster; lo cual determinó que las zonas a priorizar por tipo de amenaza, son las zonas que corresponden a las clases 2 y 4 en color amarillo y blanco correspondientemente en donde se evidencia el número de viviendas afectadas por los diferentes tipos de amenazas asociados a las inundaciones.

Cabe resaltar que para el caso de los dos mapas de las figuras 20 y 21 resultantes de la aplicación de las metodologías de Polígonos de Thiessen para priorizar por área de afectación, e Interpolación IDW ( Distancia inversa ponderada, por sus siglas en inglés) para priorizar por tipo de inundación, se pudo determinar de forma gráfica, que las zonas resultantes a priorizar se traslapan en consecuente relación con las zonas de influencia de amenaza por inundación: media y alta con las que cuenta el distrito y el IDIGER como base de zonificación de amenazas por inundación para la ciudad de Bogotá.

Figura 19. Porcentaje de viviendas afectadas por tipo de inundación.

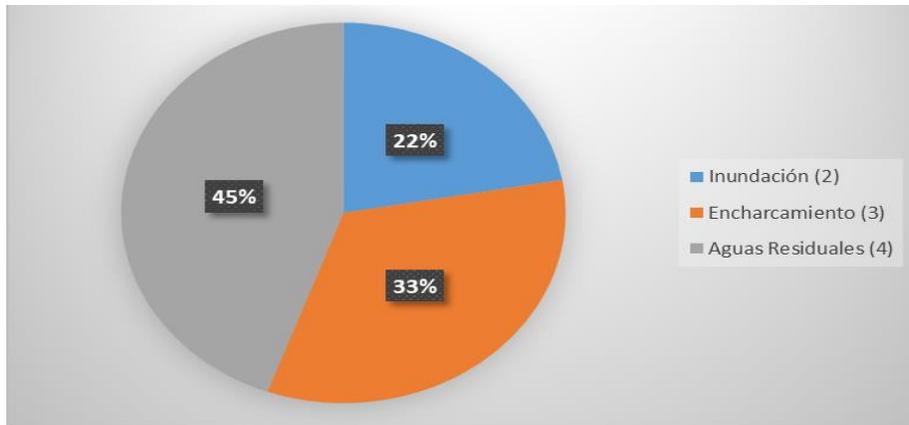


Figura 20. Aplicación polígonos de Thiessen

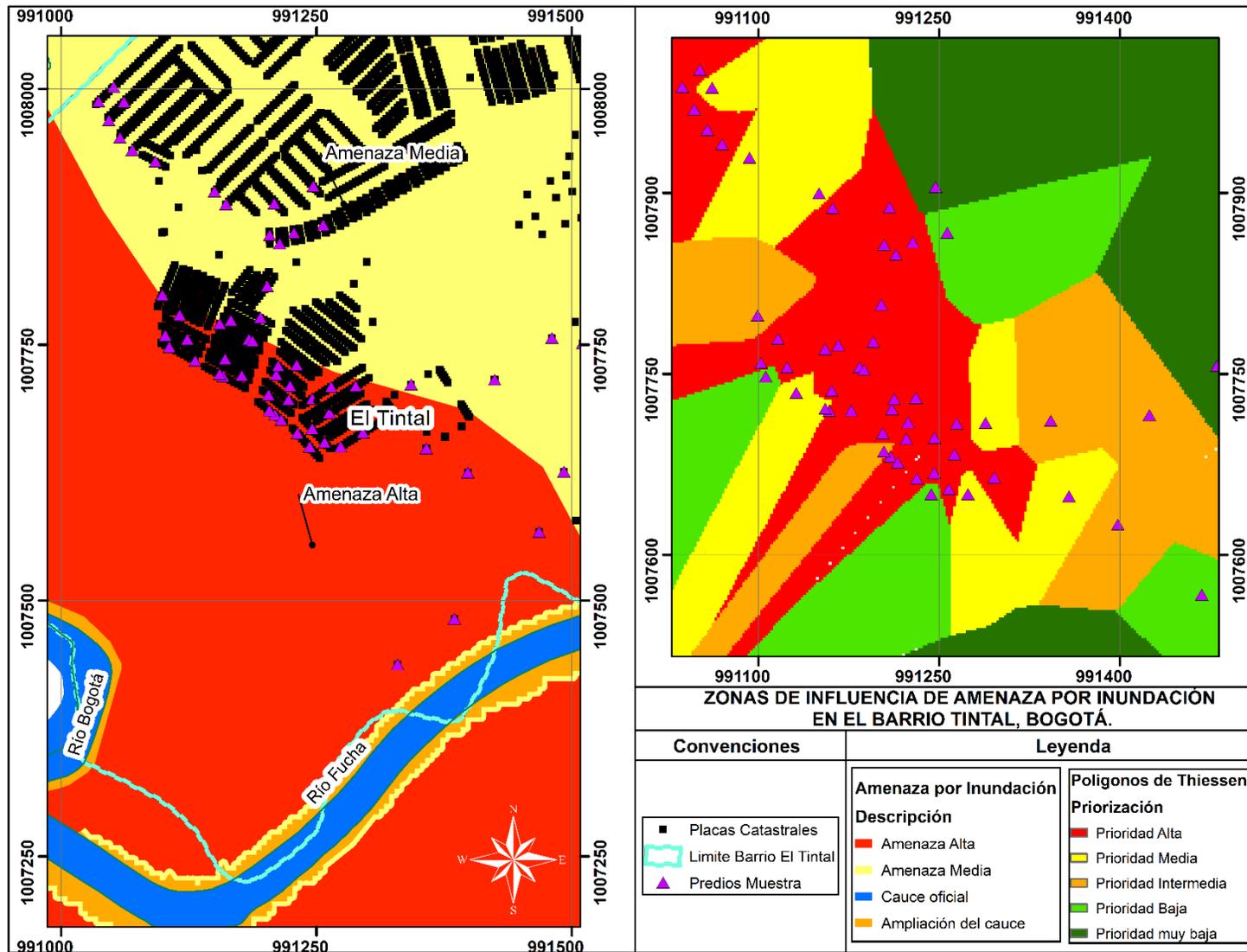
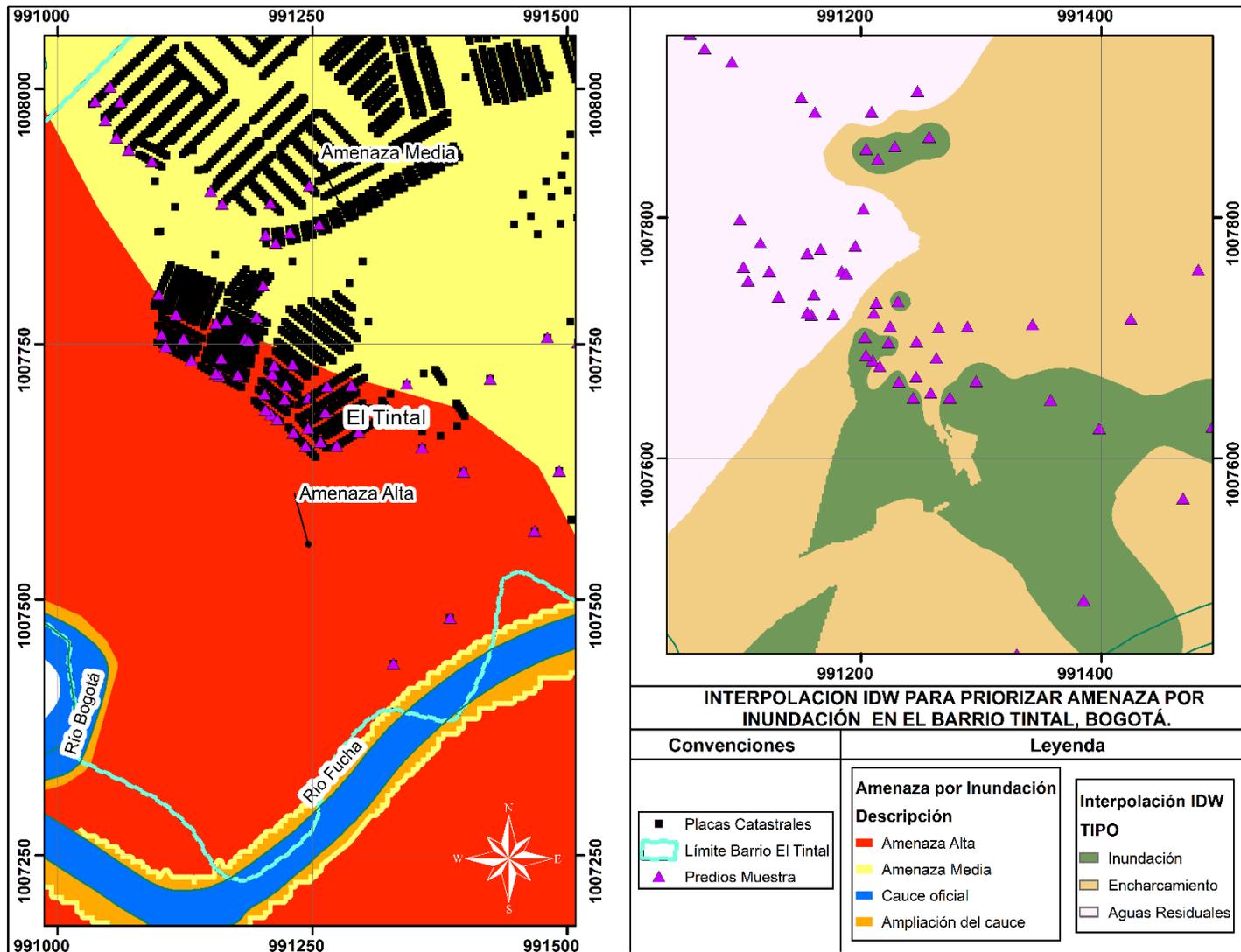


Figura 21. Aplicación interpolación IDW



La salida gráfica de los datos se puede dar de diferentes formas, offline en forma de mapa en formatos aptos para impresión, o en formato digital para ser publicado en la web, o bien generando un sistema de información geográfico en la web, que se alimente constantemente de los diferentes datos recolectados y que se encuentren disponibles para consultas en la red por parte de la ciudadanía; a continuación se presentaran las salidas gráficas de acuerdo con la norma técnica colombiana, de cada uno de los ejemplos anteriormente realizados:

Figura 22. Zonas de amenaza por inundación en Bogotá.

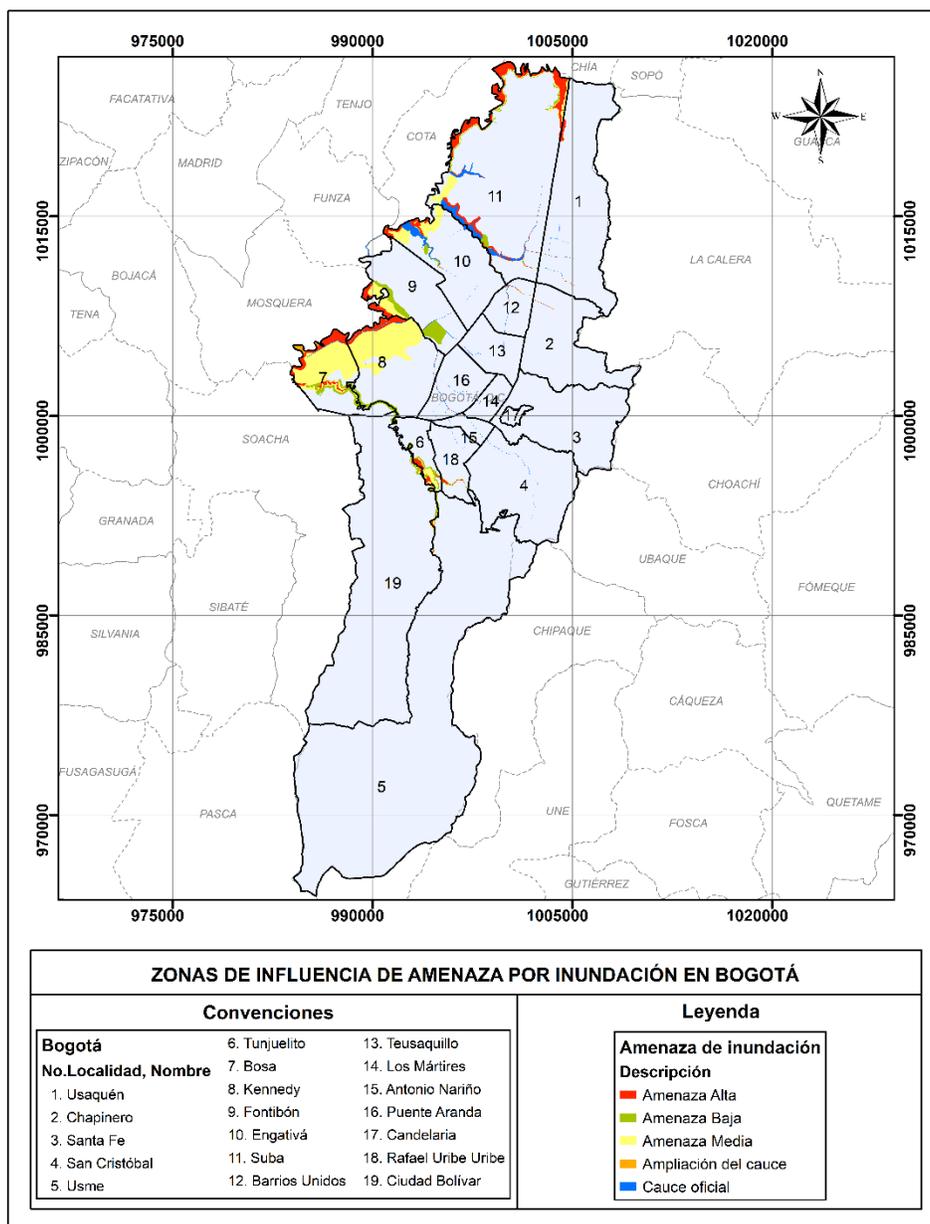


Figura 23. Zonas de amenaza por inundación en el barrio el Tintal-Bogotá.

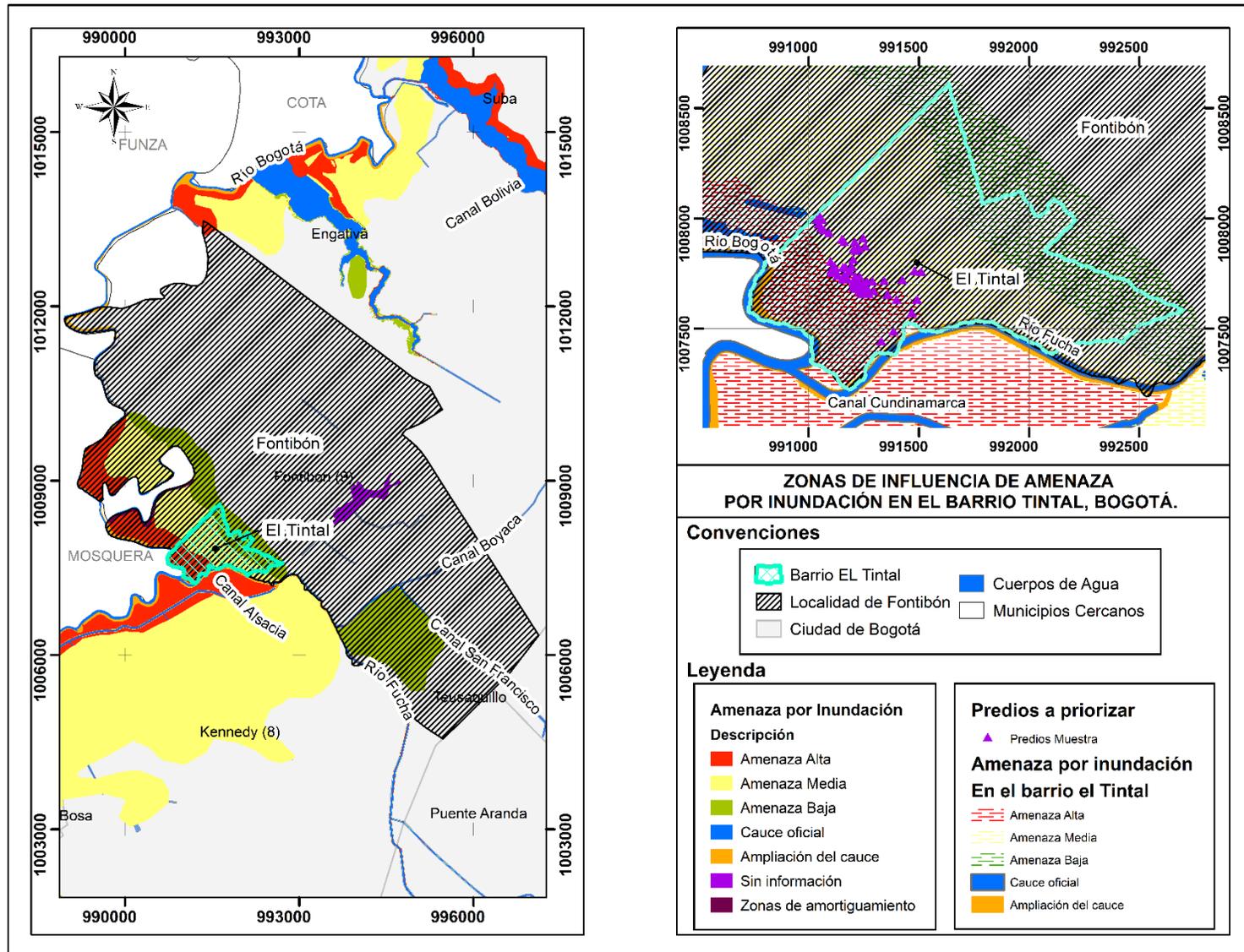


Figura 24. Aplicación de los polígonos de Thiessen.

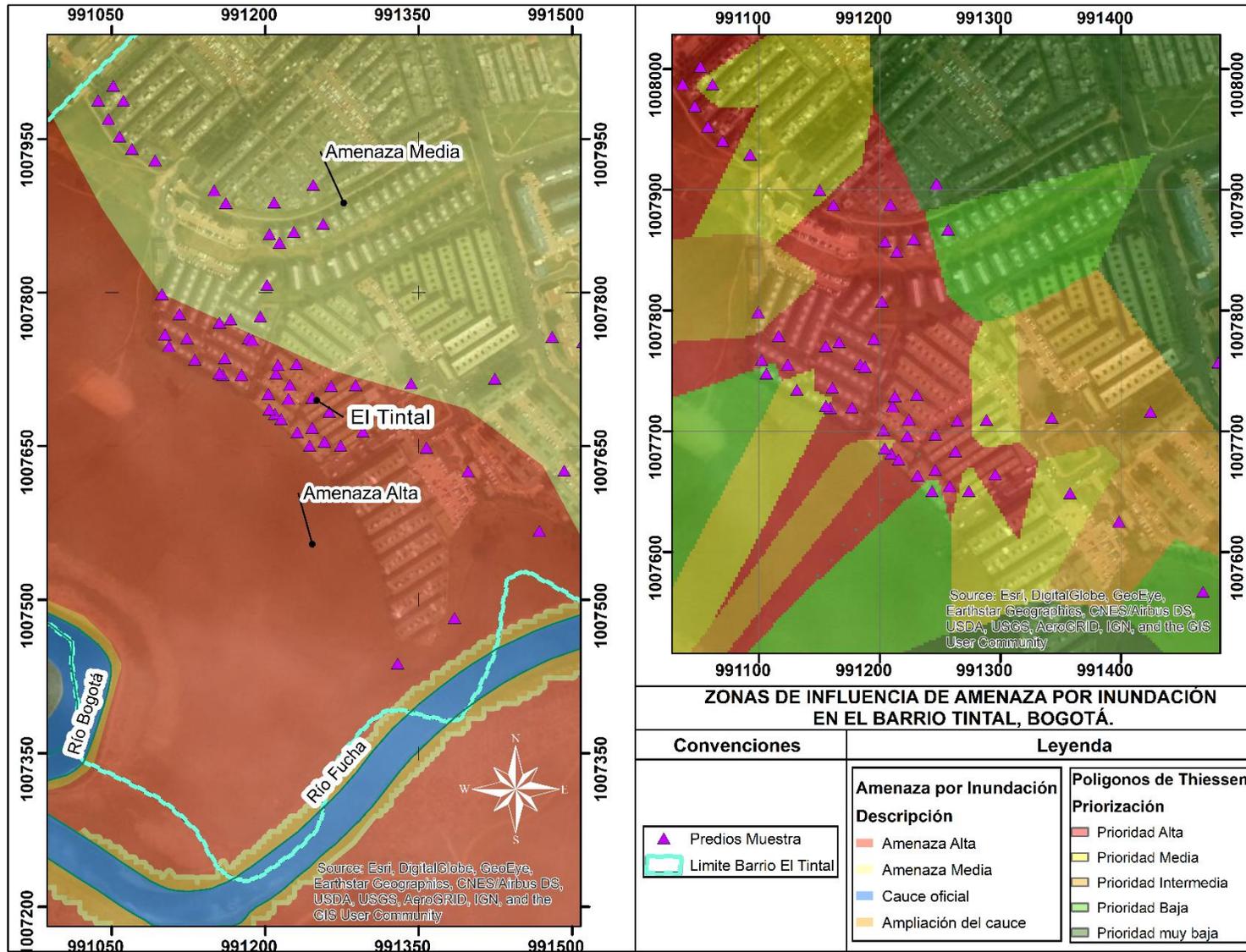
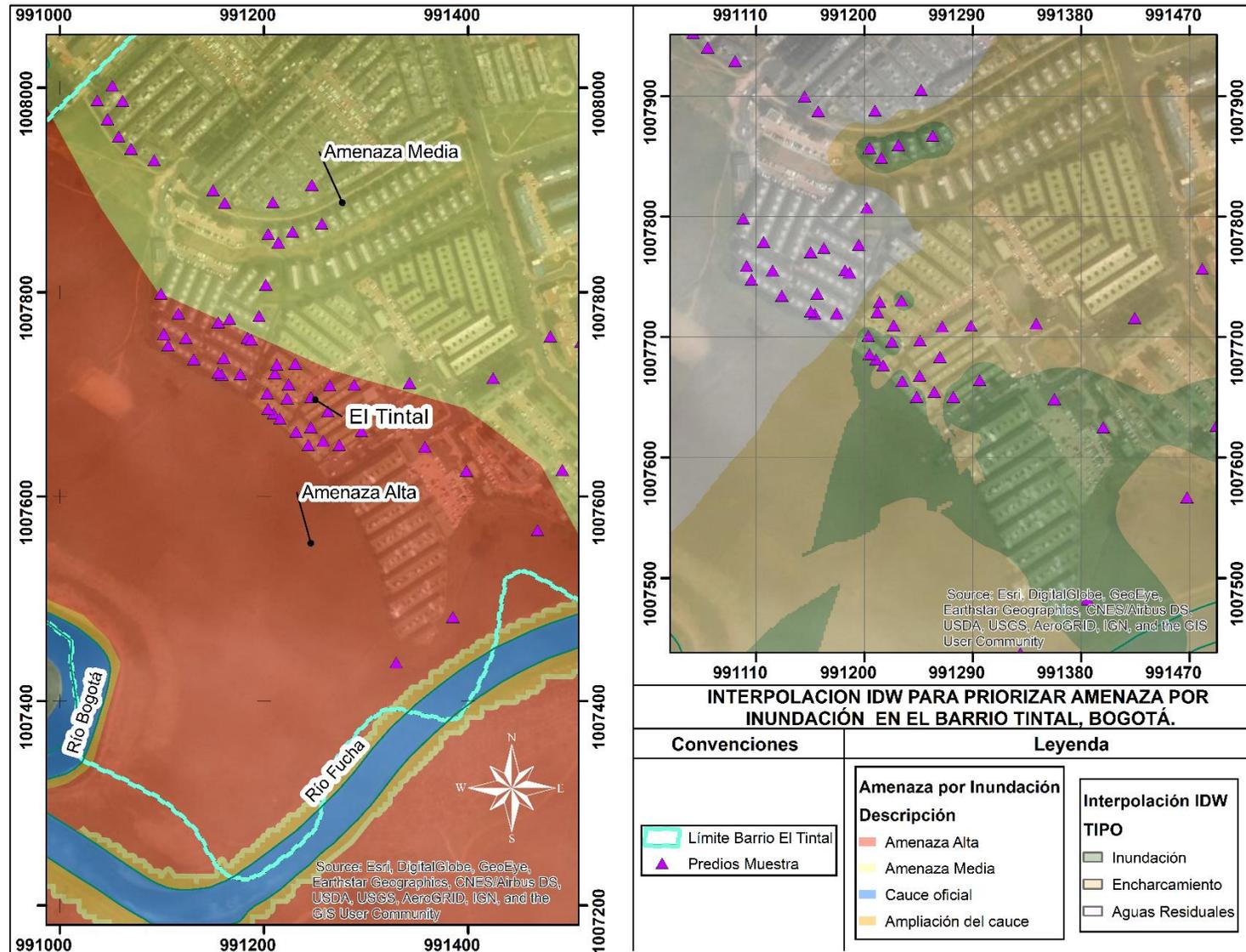


Figura 25. Aplicación de interpolación IDW



En resumen, esta guía está diseñada para ayudar a los ejecutores de los procesos de atención a desastres por inundaciones a obtener conocimientos básicos y metodológicos sobre cómo usar herramientas tecnológicas y servicios informáticos. Además, esta se centra en el uso de herramientas tecnológicas y cartográficas que pueden ser de gran utilidad en la fase de respuesta, sin incluir esfuerzos enfocados en la reducción de riesgos de desastres.

Las herramientas y procedimientos considerados para priorizar zonas de amenaza por inundación, a partir de los datos de las redes sociales, se proponen de forma concreta en las figuras 26 y 27, las cuales muestran, en primer lugar, una infografía sobre los pasos metodológicos a seguir y un diagrama de flujo que abarca los procedimientos realizados en este trabajo de investigación junto con los de la propuesta.

La infografía sintetiza lo explicado en la guía metodológica que se propuso, esta tiene en cuenta 6 pasos importantes como: la recolección de datos en Twitter, el mapeo de los mismos, la creación de una base de datos, la entrada de los datos, la gestión y la salida de datos. En cada uno de ellos se presentan los elementos y acciones que pueden ayudar a la priorización de zonas de amenaza por inundación, en una zona de estudio específica, de forma eficiente y completa, debido al mayor análisis que se puede realizar con la cartografía.

En cambio el diagrama de flujo muestra los procesos desarrollados en esta investigación junto con los pasos de la guía. Es decir que, contiene lo que se llevó a cabo con la comparación de otras metodologías aplicadas al Big Data, las visitas y entrevistas con los funcionarios de la entidad para analizar debilidades, oportunidades, fortalezas y amenazas, el ejemplo de uso del Big Data con la herramienta Eclipse y Weka; y la demostración de uso de herramientas cartográficas para la priorización de zonas.

Figura 26. Pasos metodológicos para priorizar zonas de atención frente a inundaciones mediante el uso de redes sociales

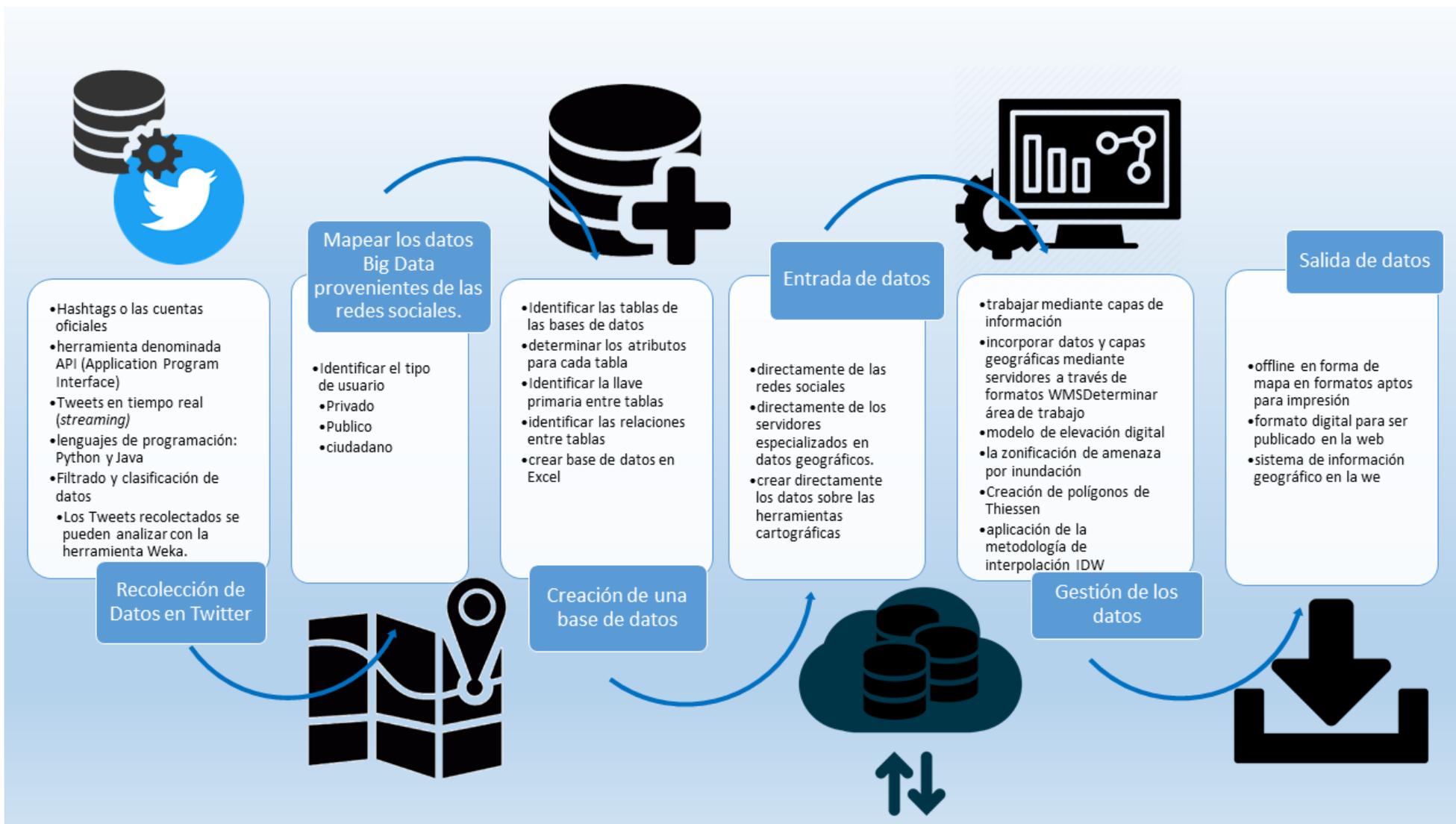
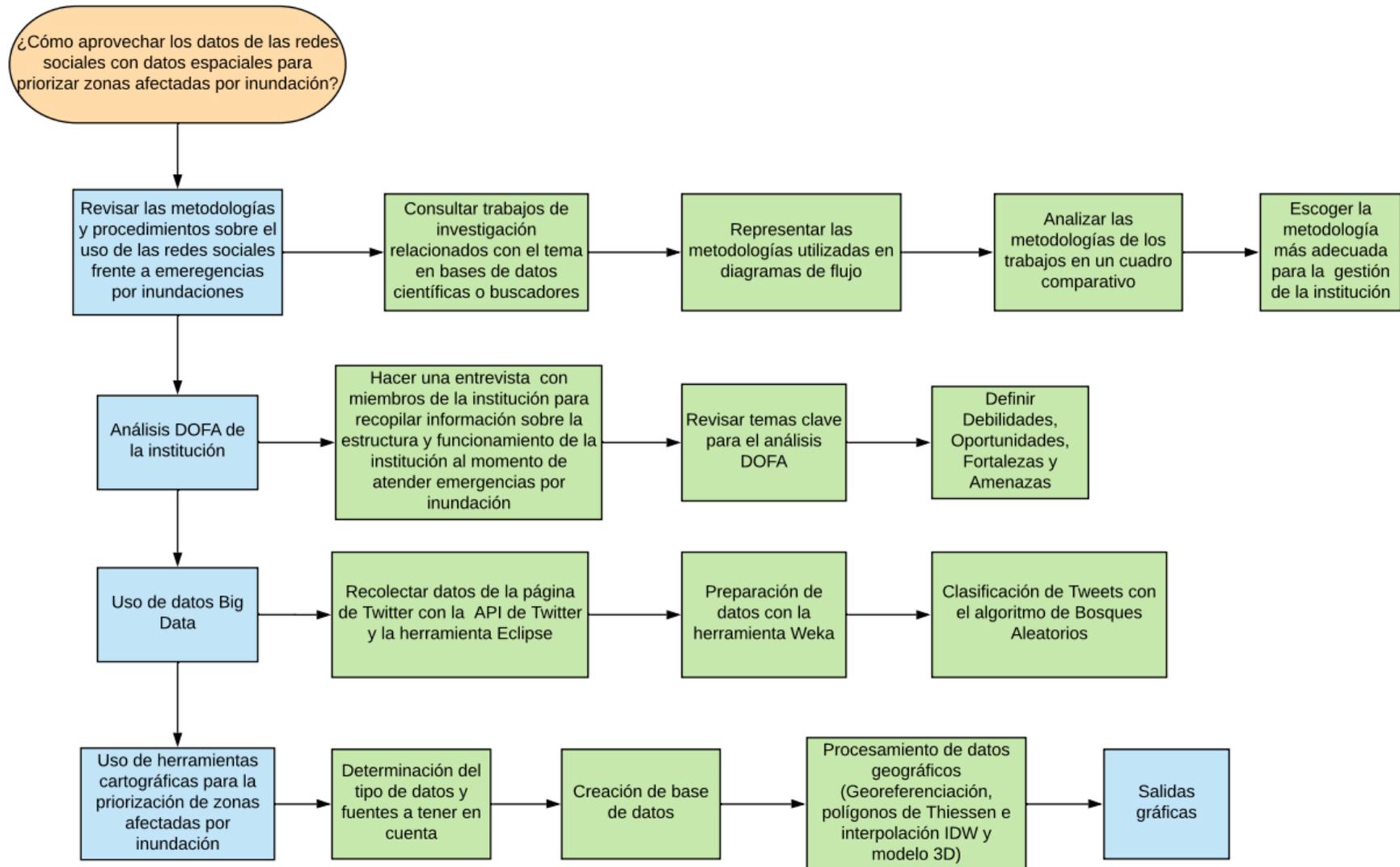


Figura 27. Diagrama de flujo general



## 6. DISCUSIÓN

Teniendo en cuenta los resultados obtenidos en la búsqueda de investigaciones relacionadas con el uso de Big Data en la atención de emergencias por inundaciones, se pueden evidenciar varias similitudes y diferencias (ver tabla 3) entre las metodologías utilizadas en cada trabajo.

Las similitudes radican en que la mayoría de las investigaciones utilizan como fuente de recolección de datos la página de Twitter, los tweets recolectados tienen al menos una palabra clave relacionada con las inundaciones y en el análisis de estos se realiza una categorización o codificación. En contraste, las diferencias se encuentran de manera general en las áreas de estudio, los objetivos, objetos de estudio, los métodos y herramientas utilizadas. Lo anterior indica grados de accesibilidad y complejidad que arrojan diferentes resultados.

Al comparar el nivel de complejidad de los trabajos, se encuentra que: Explorando el uso de las redes sociales durante la inundación de 2014 en Malasia (Aisha et al., 2015) y Dominar las redes sociales: un análisis de las comunicaciones del condado de Jefferson durante las inundaciones de Colorado 2013 (Hiltz et al., 2013), presentan un bajo nivel de complejidad con respecto a las demás metodologías consultadas porque se usan encuestas y entrevistas para contextualizar la forma en que las personas utilizan las redes sociales durante una inundación, y no se extraen la información directamente de los datos generados en las redes sociales. Sin embargo, la complejidad aumenta en el trabajo desarrollado en el condado de Jefferson porque aparte de la entrevista realizada, hay una recopilación de una gran cantidad de datos de distintas redes sociales, que posteriormente son codificados en 8 categorías.

Lo anterior, permite inferir que las entrevistas y encuestas son métodos asequibles que proporcionan un contexto inicial en todas las investigaciones, porque, en el caso de las entrevistas se obtiene profundidad y complementariedad de la información mediante el uso de una visión propia y ajena, en la que se puede averiguar por opiniones relacionadas a un evento (Díaz-Bravo, Torruco-García, Martínez-Hernández, & Varela-Ruiz, 2013); y en las encuestas, se recolectan y analizan datos de una muestra representativa de una población con el fin de obtener medidas de los conceptos derivados de una problemática. (López-Roldán & Fachelli, 2016).

Además si se contrastan los resultados de estas dos investigaciones, que usaron los métodos descritos anteriormente, la encuesta aplicada en Malasia (2014) muestra que Facebook es la red social más usada en comparación a los blogs, Instagram, o Twitter. Asimismo se encontró que durante el periodo de inundación se usan otro tipo de tecnologías de la comunicación como WhatsApp, llamadas y SMS. En cuanto al trabajo del condado de Jefferson se encontró que Twitter es la red social más efectiva para que la información llegue rápidamente a una amplia audiencia, mediante la creación de hipervínculos que dan a conocer el blog del equipo, la página de Facebook y el mapa como fuentes de información valiosas.

Es necesario resaltar que la investigación que más se relaciona con el tema de este estudio es la investigación del condado de Jefferson debido a que tiene como objeto de estudio al Equipo de Gestión de Incidentes Tipo III del condado de Jefferson, el cual es el encargado de responder a situaciones de gran peligro en el Condado de Jefferson y las áreas circundantes. Su equipo está formado por 20 agencias diferentes dentro y alrededor del condado, incluidas las agencias de cumplimiento de la ley, departamentos de bomberos, servicios forestales y de administración de tierras (Sheriff's Office Jefferson County).

La subdirección de manejo de emergencias y desastres del IDIGER, la cual tiene como función principal liderar, organizar e integrar el proceso de manejo de emergencias y desastres para que la respuesta a emergencias y recuperación sea la más efectiva (IDIGER, 2018). Cabe aclarar que ambas investigaciones se enfocan en instituciones encargadas de atender emergencias por inundaciones, sin embargo la metodología del trabajo en Jefferson no explica de forma clara las herramientas o técnicas utilizadas para analizar los datos de las redes sociales, así como la investigación de Malasia, que solo considera las mediciones de las encuestas y no una aplicación del Big Data en la que se procese una gran cantidad de datos; lo que quiere decir que no es viable tener como base estas metodologías en la propuesta de esta investigación.

A pesar de lo descrito anteriormente, existe una tendencia a utilizar Twitter, la cual se encuentra en el resto de investigaciones consultadas como: Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres (Nair et al., 2017), Chatter on The Red: ¿Qué revela

la amenaza sobre la vida social de la información generada en microblogs? (Starbird et al., 2010) Y Big Data Analytics para la respuesta ante desastres y recuperación a través del análisis de sentimientos (Ragini, Anand, & Bhaskar, 2018b). En estos trabajos se emplea Twitter para recolectar y analizar los datos sobre las inundaciones ocurridas en un área de estudio específica, pero cada trabajo tiene una dificultad de entendimiento diferente por las herramientas utilizadas y los procedimientos ejecutados.

Es decir, que en los tres trabajos se habla de una recolección de datos utilizando Twitter, en el caso de Chennai 2015 se recolectaron únicamente los tweets que contenían la palabra #chennaiflood y se relacionaban con la inundación en el área de Tamil Nadu, Chennai. La metodología es de fácil entendimiento porque los tweets se clasificaron con la herramienta gratuita Weka a través de 3 algoritmos de aprendizaje automático: Árboles de Decisión, Bosques Aleatorios y Naive Bayes, obteniendo como categorías resultantes: Necesito Ayuda, Medidas de Alivio, Expresar Gratitud, Quejas y Otro.

De igual forma se pudo observar que los usuarios más influyentes fueron tanto organizaciones gubernamentales como individuos y que el mejor algoritmo para una clasificación correcta es Bosques Aleatorios, porque tiene menos errores en la clasificación. Este algoritmo selecciona de forma aleatoria un número de variables con los que se crean los árboles individuales, a partir de ellos se hacen predicciones de la clase más votada para obtener un resultado de clasificación. (Merino & Chacón, 2017) Permite trabajar con una gran cantidad de variables y reconocer las más significativas, es práctico en la estimación de datos faltantes, manteniendo su precisión, y la preparación de los datos es mínima. (Alvear, 2018).

En el trabajo del Río Rojo se recolectaron los datos por la API de búsqueda de Twitter, teniendo en cuenta las palabras claves red river y redriver y que el río Rojo fluye desde el sur de Fargo hacia el norte a lo largo de Dakota del Norte (ND): frontera estatal de Minnesota (MN) en los EE. UU. Y en el Lago Winnipeg, justo al norte de la ciudad de Winnipeg en Manitoba, Canadá. Este procedimiento es de mayor complejidad debido a que se analiza una gran cantidad de tweets, se visualizan en tiempo real con un programa llamado E- Data Viewer y se codifican cualitativamente.

Es preciso señalar que este software analiza grandes conjuntos de datos basados en CMC, esto comprende una gran cantidad de posibilidades comunicativas dadas a través del uso de una computadora conectada a la Red para crear, almacenar, transferir y procesar información (Rodríguez, Corrales, & Rodríguez, 2016); y que el software E Data Viewer es únicamente para uso interno y permite visualizar los datos de CMC a lo largo del tiempo, esto indica que no es viable tener en cuenta esta investigación para la propuesta por el acceso limitado a la herramienta.

La investigación hecha en 2018 sobre el análisis de sentimientos fue realizada en varios países asiáticos, la recolección de datos se hizo en Twitter buscando tweets con las palabras clave: HUDHUD, Vizag flood, Nilofar, Kashmir floods, India-Pakistan floods y Pakistan floods. La metodología de este trabajo es de mayor complejidad porque al ser un análisis de sentimientos en el que se extrae, de forma automática y con técnicas computacionales específicas, la información subjetiva de un texto sobre un tema determinado presenta una dificultad en la interpretación de resultados porque aún no se sabe si tiene una connotación negativa o positiva.

En otros términos, es difícil ejecutar correcta y efectivamente los procesos implicados en el análisis de sentimientos, además las palabras que expresan algún sentimiento dependen del contexto en el que se encuentren. Por ejemplo, no es lo mismo decir “Tiene la cabeza bien puesta” a decir “Está perdiendo la cabeza”, la primera frase tiene un significado positivo mientras que la otra uno negativo. También hay dificultades en el manejo de la negación en las oraciones, debido a que hay negaciones que no afectan el sentido de la oración, tal como: “No hay duda de que es el mejor”, pero si hay negaciones que cambian totalmente el sentido de la frase como: “La película no es buena.” (Sande, 2018).

Hay más situaciones en las que el análisis de sentimientos es complicado de usar como la ambigüedad en las oraciones, la interpretación de la ironía o sarcasmo y otros. Estas son razones por las que se descarta el uso del análisis de sentimientos en la propuesta metodológica de este proyecto.

Dentro de la investigación de Inundación a través del lente de la actividad del teléfono móvil se estudió la viabilidad de usar los registros de llamadas (CDR) para interpretar las inundaciones ocurridas en Tabasco 2009. En primer lugar, se usaron los datos del censo (2010) y de áreas administrativas globales para evaluar si había una gran representación de los usuarios de llamadas, luego se utilizaron herramientas Python para analizar los datos de CDR y de teledetección. Puntualmente lo que se hizo fue filtrar los datos de las llamadas según las coordenadas GPS asociadas, recrear trayectorias de desplazamiento y finalmente visualizar los datos geo-referenciados.

Esta metodología de trabajo se tuvo en cuenta para la propuesta debido a que integra el componente geográfico con el manejo de una gran cantidad de datos. A pesar de que se trabajan con llamadas, se tiene gran énfasis en la mapificación de los datos, lo cual se vio como una dificultad del IDIGER a la hora de atender emergencias por inundación. Se tiene conocimiento de que el IDIGER aun maneja las llamadas mediante el NUSE, pero si se hiciera un procesamiento de imágenes satelitales como en este trabajo para ver el área inundada, una comparación con datos de catastro, los puntos de las coordenadas de las llamadas y el análisis de tweets en tiempo real, tendría como resultado la complementariedad de la información, la reducción de intermediarios, la priorización, rapidez y eficacia en la atención a emergencias.

Es por esto que la metodología propuesta en este trabajo de investigación se basó en el trabajo: “Uso y análisis de Twitter durante la Inundación de Chennai (2015) para la gestión de desastres”, porque en la parte de uso del Big Data la metodología es de fácil entendimiento, hay un fácil acceso a la herramienta y comprueba que hay algoritmos más apropiados para clasificar los datos recolectados de Twitter. También se eligió el trabajo de las llamadas para tener un componente geográfico en la propuesta que permita complementar el uso del Big Data en la atención de emergencias por inundaciones con un mayor análisis de variables geográficas.

Mediante el análisis realizado a la subdirección del IDIGER partiendo de las necesidades encontradas, así como de las oportunidades de aplicación y replicación de metodologías de aplicación del uso del Big Data en otros países dentro de las respuestas institucionales en la atención a emergencias causadas por inundaciones, que dio como resultado la creación de una

matriz DOFA; se encontraron una serie de estrategias que se consideran dentro de la transversalidad fortalezas oportunidades, debilidades oportunidades, fortalezas y amenazas, y debilidades amenazas; la obtención de información e insumos de diferentes entidades para alimentar la cartografía es fundamental, ya que al tener fácil apoyo de las demás entidades, también se puede lograr una fácil participación y colaboración activa por parte de las mismas, que permita generar una completa estandarización de los datos, reduciendo así de manera significativa los tiempos de procesamiento y análisis de la información proveniente de los datos, también, la generación de rutas de atención efectivas mediante el aprovechamiento de la georreferenciación, es una de las mayores oportunidades a tener en cuenta apoyado en la ayuda de medios digitales y sistemas de información geográfica.

Otra de las oportunidades que se pueden evidenciar es la reducción de los tiempos de capacitación a personal frente a una emergencia, ya que como se nombró anteriormente no existe una jerarquía entre entidades, y se debe capacitar a todo el personal disponible para lograr una atención integral, lo que se puede lograr mediante la continua capacitación del personal y la reducción de rotación del mismo por parte de las entidades.

La creación de una línea de atención directa, es una de las formas más efectivas con las que contaría el IDIGER, para ser una entidad más cercana al ciudadano conociendo de primera mano las necesidades del mismo, obteniendo directamente de la ciudadanía información primaria necesaria para la creación y mantenimiento de mapas de calor de focos de inundaciones en Bogotá, presentando así una oportunidad de generación de cartografía propia y actualizada, la cual estaría a disposición de la entidad en cualquier momento, para llevar a cabo análisis espaciales de cualquier nivel y de cualquier fenómeno.

No obstante, se debe tener en cuenta que se deben integrar los componentes X e Y de las bases de datos, pues como se ha venido hablando el componente geográfico es lo que le da el plus a las empresas o entidades, como es el caso de Google, Amazon, Facebook entre otros, que han explotado la relación entre el componente geográfico y el almacenamiento de estos mismos datos con el fin de determinar patrones que facilitan la aplicación e interpretación de algoritmos en

correlaciones de datos para predecir necesidades, y luego aplicar dicho conocimiento de necesidades para generar productos y servicios.

En el caso del IDIGER la aplicación del conocimiento y análisis de la información se empleara para el mejoramiento de los tiempos de atención, otro factor a considerar es la adición de un componente Z en el conjunto de datos geográficos, pues en grandes ciudades donde la propiedad horizontal es predominante se requiere saber, desde qué piso se está generando la información y así también aprovechar la información de las redes sociales.

Incrementar y mejorar la capacidad de organización y coordinación entre las entidades y el IDIGER, es uno de los factores a mejorar, pues las demás entidades son externas al IDIGER, no existe una jerarquía por lo tanto se pueden generar retrasos en los tiempos de atención, mejorar la coordinación y cómo se gestiona es fundamental, así mismo mantener informada y capacitada a la comunidad frente al manejo y utilización de los nuevos recursos virtuales como la creación de un “Hashtag” o demás comportamientos comprendidos dentro de la huella digital de dichas comunidades con el fin de aprovechar la información generada en las redes sociales.

La propuesta de uso de la información obtenida de los datos generados en las redes sociales para la priorización de zonas afectadas, así como el apoyo en la toma de decisiones dentro de las respuestas puede resultar efectiva porque se utilizan herramientas y procedimientos que se adecuan a las características de los datos. Una de estas herramientas es Eclipse, que se puede descargar e instalar fácilmente de forma gratuita, permitiendo la construcción de un entorno de desarrollo mediante la conexión de componentes plug ins. Para el desarrollo de la propuesta de este trabajo este programa fue efectivo en la recolección de Tweets porque se almacenaron en tiempo real los Tweets de las cuentas y hashtags relacionados con inundaciones y emergencias.

También el programa Weka es gratuito, tiene gran variedad de herramientas para transformar los datos, actividades de clasificación, regresión, clustering, asociación y visualización; y está orientada a la extensibilidad por lo que se pueden construir nuevas funcionalidades. A pesar de todas estas funcionalidades, se requiere de un científico de datos que pueda configurar todas las opciones necesarias para preparar los datos y posteriormente clasificar usando el algoritmo de

Random Forests. Este algoritmo es el más preciso porque al combinar árboles de decisión hace que existan más opciones para clasificar el valor de entrada y poder elegir la clase más popular.

En la generación de zonas de priorización mediante representación cartográfica se obtuvo que la utilización de polígonos de Thiessen dio buenos resultados para analizar diferentes zonas de afectación por área, y se pudo determinar también que las zonas óptimas para priorizar son las zonas de menor área que se encuentren cerca la una de la otra y que en conjunto generen una gran zona que da como resultado un área específica con mayor densidad de personas afectadas por metro cuadrado; también se encontró que una de las mejores formas para priorizar es mediante el proceso de interpolación de distancia ponderada inversa o IDW que a diferencia del proceso anterior identifica y prioriza las zonas no por área, sino por tipo de emergencia asociado a inundaciones para lo cual es necesario asignar un único número que identifique el tipo de amenaza; ambos resultados dieron una similitud considerable entre ellos en cuanto a las zonas óptimas para priorizar, así mismo dichas zonas se encuentran ubicadas en las zonas correspondientes a amenaza alta por inundación, lo cual se pudo notar que es un modelo correspondiente que puede acercarse a la realidad, y ayudar a priorizar áreas de atención frente a una emergencia por inundación.

En cuanto a la ubicación de las estaciones meteorológicas como resultado de la aplicación de los geo procesos necesarios para llevar a cabo un análisis multicriterio, se obtuvieron 45 zonas estratégicas con las condiciones óptimas, teniendo en cuenta los criterios que se plantearon durante el desarrollo de este trabajo, así mismo se utilizaron dos capas temáticas de datos climáticos relevantes como lo son la temperatura y la precipitación, con el fin de comparar y enfatizar, si, las zonas resultantes, son las más adecuadas para el proceso de instalación de una estación meteorológica de cualquier tipo.

Las zonas más adecuadas para la instalación de estaciones meteorológicas son áreas que tengan un clima templado para que los datos medidos sean homogéneos. La zona que se obtuvo como óptima fue una zona muy pequeña debido a que las demás estaciones se encuentran sobre climas fríos y cálidos. Además, en la localidad del Sumapaz hay estaciones que se localizan solo cumpliendo los criterios de cobertura, pendiente, cuerpos de agua, vías y estaciones preexistentes. Falta información de temperatura y la generación de las nuevas estaciones puede ser una solución para

esta problemática. Los procesos necesarios para llevar a cabo el análisis multicriterio de las zonas óptimas para instalar estaciones meteorológicas como red de apoyo y monitoreo para la subdirección, puede verse reflejado en el anexo 6.

## CONCLUSIONES

- La comparación de trabajos aplicados en otros lugares y relacionados con el uso de Big Data en la atención a emergencias por desastres tienen diferentes similitudes y diferencias que se pueden apreciar en los procedimientos realizados y en las herramientas empleadas, esto representa cierto nivel de complejidad para la propuesta presentada en este proyecto por el acceso a las herramientas y la comprensión de las técnicas y procesos usados. Se encontró de forma general que en los trabajos hay similitudes en los pasos para preparar y analizar los datos como: La recolección de datos, filtrado y clasificación de datos; también se vio que la red social más usada fue Twitter. Las diferencias se observaron en los objetivos, objetos, áreas de estudio y en las diversas técnicas para clasificar los datos.
- El sistema de llamadas actual del IDIGER es útil para atender emergencias, pero la aplicación del Big Data con el análisis de redes sociales y la representación cartográfica del área inundada, datos de catastro, los puntos de las coordenadas de las comunicaciones y el procesamiento de imágenes satelitales, pueden ayudar a que la atención durante una emergencia por inundación se base en información veraz y real para priorizar en envío de ayudas requeridas por los afectados en una localización conocida.
- Hay posibilidad de que la metodología propuesta en esta investigación pueda ser aplicada por entidades gubernamentales encargadas de atender emergencias si se realiza una revisión más profunda de todas las configuraciones que se pueden realizar para preparar los datos de Twitter y clasificarlos con un algoritmo que tenga buena precisión para que las predicciones sean acertadas durante una emergencia. Lo anterior requiere de un científico de datos o un profesional que esté capacitado para manejar lenguajes de programación propios de estos procedimientos.
- En la actualidad un SIG es la herramienta que mejor permite combinar e integrar gran cantidad de información variada que no es fácilmente interpretable por otros medios, ya que este combina información proveniente de cartografía base, con otra u otras capas de información de cualquier tipo, ya sea estadística, matemática, temática entre otras haciendo

de esta manera que las consultas sobre el territorio se puedan realizar de manera rápida y que puedan aportar información necesaria, versátil y mucho más rica para el usuario.

- El manejo adecuado de la información mediante la gestión de bases de datos electrónicas permitirá el fácil procesamiento y análisis de la información cartografía de la entidad como tal; y por otro lado logrará estandarizar la mayoría de la información que estará disponible para cualquier entidad.
- Integrar los componentes espaciales X e Y es uno de los factores más importantes para lograr una adecuada georreferenciación y posterior representación cartográfica de la información, que ayudara a determinar y priorizar las zonas de atención en pro de optimizar la gestión de la atención a emergencias por inundación.
- Se debe considerar también la integración del componente espacial Z de altura en la recolección de la información que alimenta las bases de datos que a su vez sirven para alimentar la cartografía, ya que en ciudades densamente pobladas y en crecimiento como Bogotá, cuyo régimen de crecimiento predomina la propiedad horizontal, es necesario saber desde que altura se está generando la información, lo que enriquecerá el análisis espacial de las zonas afectadas.
- Existe un desaprovechamiento de la información obtenida de los datos generados en las redes sociales, dentro de los procesos de atención a emergencias por inundaciones, debido a que existe desconocimiento de herramientas, poca normatividad y falta de integración entre las entidades oficiales y las comunidades; la adecuada representación de los datos mediante el uso de cartografía en donde se presenten las zonas priorizadas, es una herramienta que genera valor agregado a la información extraída de los datos, ya que de esta forma es más fácil visualizar los datos contribuyendo de manera significativa a los procesos de toma de decisiones y por ende a la mejora de las respuestas institucionales; entendidas como los tiempos de atención y gestión de los recursos.

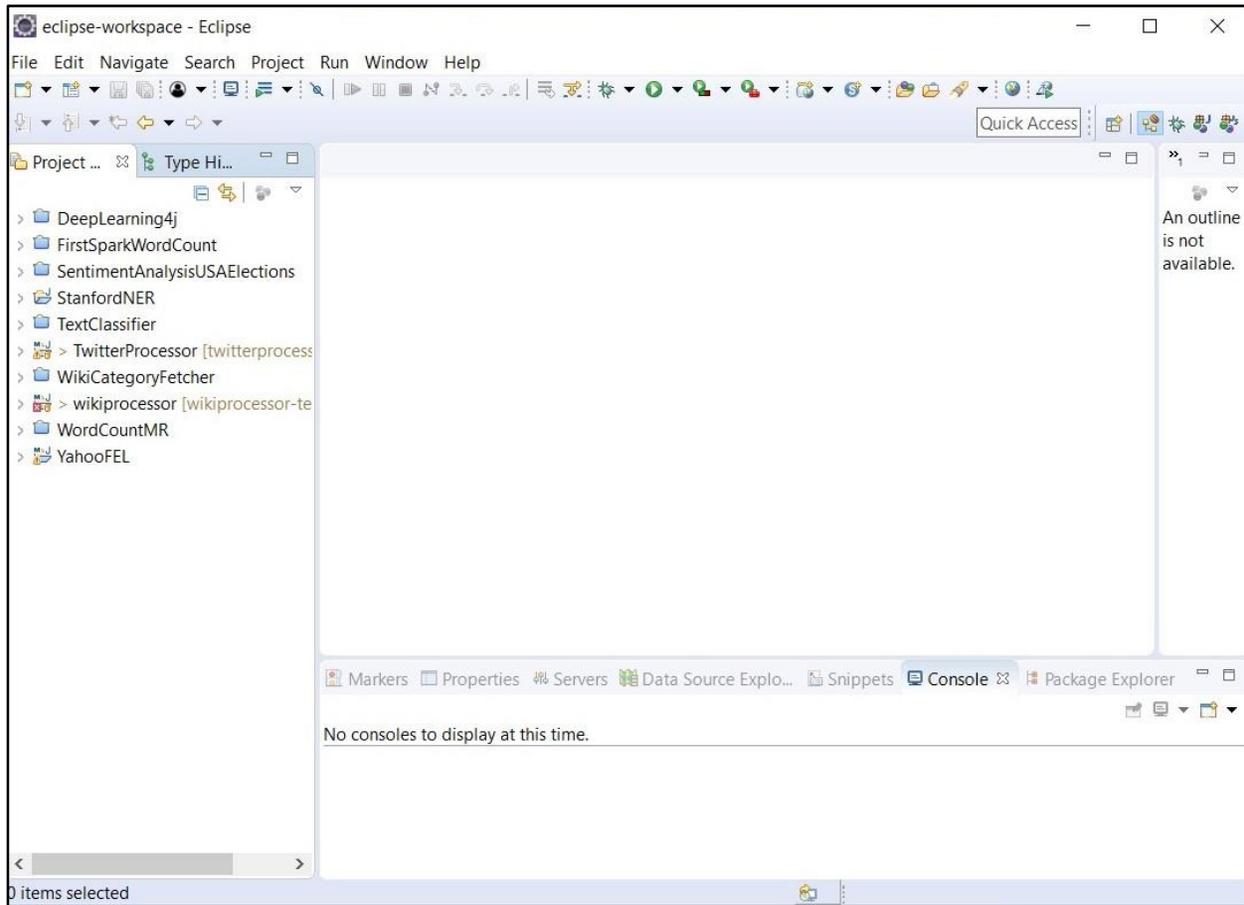
- Los datos generados en las redes sociales proveen, a través de su adecuado procesamiento y análisis, una fuente de información primordial que facilita la priorización de zonas de atención, en el caso de un desastre ocurrido por inundación, porque se conocen las coordenadas de los puntos desde los que se transmiten los Tweets, las necesidades de cada persona y su comportamiento en tiempo real. Si a lo anterior se le suma el componente geográfico con el empleo de imágenes satelitales, datos de catastro y mapas temáticos, es posible visualizar y ubicar las áreas de mayor afectación que deben ser atendidas de forma rápida y eficaz, esto último contribuye a la mejora de los procedimientos de respuestas institucionales.

## RECOMENDACIONES

- Es importante estudiar más a fondo las técnicas y herramientas de Eclipse y Weka con un profesional especializado o un científico de datos, quien es capaz de evaluar distintas fuentes de información de cualquier organización, extraer datos en diversos formatos, analizarlos, desarrollar los algoritmos adecuados, inferir, preparar y explicar los resultados de dicho análisis. Además puede probar con otras herramientas aún más apropiadas y convenientes para usar el Big Data en la atención a emergencias causadas por inundación.
- Se sugiere una actualización de la información presentada en los mapas del IDIGER ya que se encontraron mapas de los años 2013 al 2018, esto con el fin de que la información cartográfica sea veraz, vigente y pueda ser complementada con la información producida en tiempo real, a través del uso de técnicas y procedimientos de análisis de una gran cantidad de datos provenientes de redes sociales como Twitter, para aumentar los aciertos en la toma de decisiones durante una emergencia causada por inundación.
- Es importante tener en cuenta la capacitación continua de la comunidad como fuente principal de información, para incrementar la participación activa de la misma en redes sociales, mediante propagandas y actividades educativas, ya que, sin la participación y la huella digital de la misma, no existirían datos suficientes para alimentar las bases de datos y por consiguiente la cartografía.

# ANEXOS

## Anexo 1. Entorno de trabajo de Eclipse



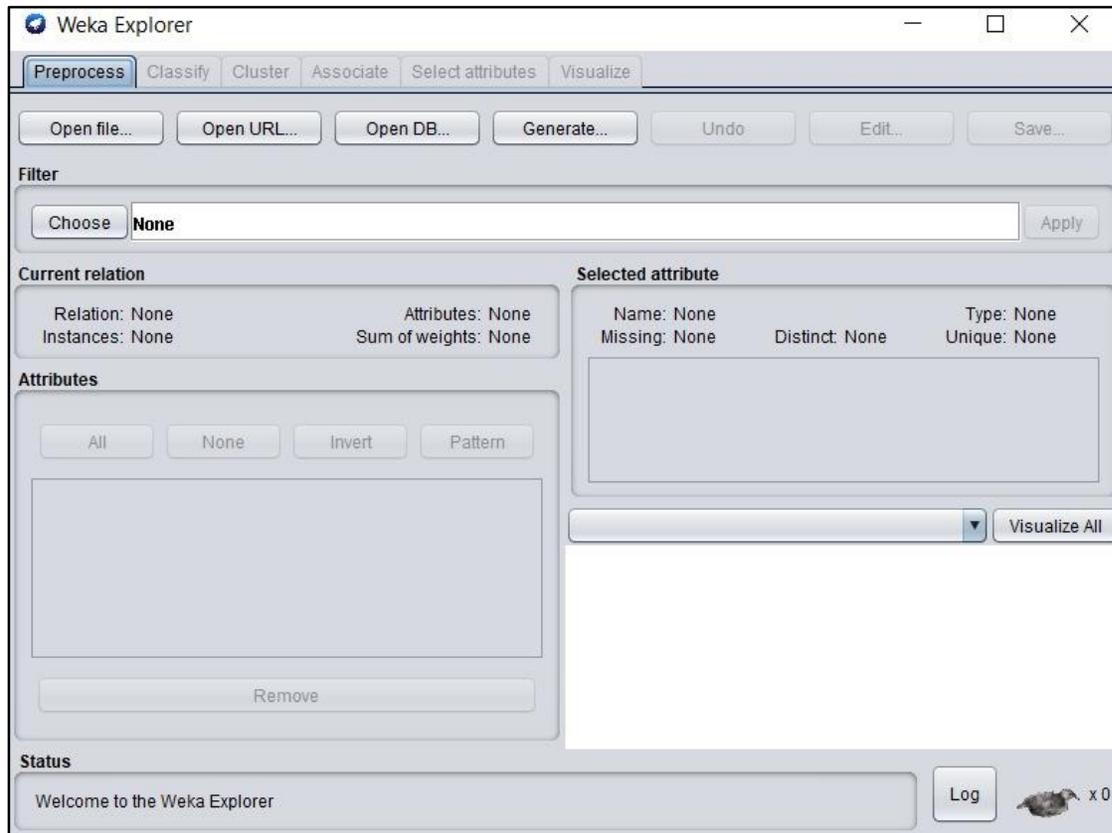
## Anexo 2. Lenguaje para definir variables y parámetros de conexión

```
public class Main {
    static FileOutputStream fos;
    1 final static String coleccion = "twitterdb";
    static DBCollection coll;

    public static void main(String[] args) throws UnknownHostException {
        ConfigurationBuilder cb = new ConfigurationBuilder();

        cb.setDebugEnabled(true).setOAuthConsumerKey("9ESH8DSaixhMEEnpGa2Dg5oAC")
            .setOAuthConsumerSecret("Pi7GbgzbjzBuaJLugsqMre13hI8RB24vXME8ISfIARmR0nctXg")
            .setOAuthAccessToken("53967924-JH0t0l0r50wPkYTJSvBg3WfJTDbLwFcEQwHbj1fc")
            2 .setOAuthAccessTokenSecret("Of5XWcz4toCkb5EKdrpV77fuJaSmo4E1P4FxKoTuMxBYN").setJSONStoreEnabled(true)
            .setIncludeEntitiesEnabled(true);
    }
}
```

### Anexo 3. Pestaña de preprocesamiento de Weka



### Anexo 4. Partes del archivo ARFF

```
1 @relation weather
2
3 @attribute outlook {sunny, overcast, rainy}
4 @attribute temperature numeric
5 @attribute humidity numeric
6 @attribute windy {TRUE, FALSE}
7 @attribute play {yes, no}
8
9 @data
10 sunny,85,85,FALSE,no
11 sunny,80,90,TRUE,no
12 overcast,83,86,FALSE,yes
13 rainy,70,96,FALSE,yes
14 rainy,68,80,FALSE,yes
15 rainy,65,70,TRUE,no
```

Tomado de: <https://www.adictosaltrabajo.com/2019/01/17/introduccion-al-machine-learning-con-weka/>

## Anexo 5. Fuente de los datos requeridos en la priorización

Las informaciones se pueden recolectar directamente de los datos de las redes sociales mediante las herramientas ya mencionadas, o bien directamente de los servidores especializados en datos geográficos, así mismo cabe resaltar que también se pueden crear directamente los datos sobre las herramientas cartográficas.

<b>Materiales</b>			
<b>Datos e información</b>			
<b>Dato</b>	<b>Fuente</b>	<b>Formato</b>	<b>Fecha/Escala/ Otros</b>
<b>1.</b> Modelo de elevación digital (DEM)	USGS <a href="https://earthexplorer.usgs.gov/">https://earthexplorer.usgs.gov/</a>	Raster dataset	Escala: 1: 25.000 Resolución: 30m x 30m
<b>2.</b> Shapefile de red vial en Bogotá	IDECA	Shapefile (.shp)	Año de Actualización: 2013. Tipo: Línea Escala: 1: 25.000
<b>3.</b> Shapefile de cuerpos de agua en la ciudad de Bogotá.	IDECA	Shapefile(.shp)	Año de Actualización: 2015-2017. Tipo: Polígono Escala: 1: 2.000
<b>4.</b> Imagen: Mapa de zonificación de zonas de amenaza por inundaciones en Bogotá	IDECA	Imagen (MAPA)	Escala:

5. Shapefile de placas catastrales de la ciudad de Bogotá.	IDECA	Shapefile(.shp)	Año de Actualización: 2016-2017. Tipo: Punto Escala: 1: 2.000
6. Shapefile de barrios catastrales de la ciudad de Bogotá.	IDECA	Shapefile(.shp)	Año de Actualización: 2015-2017. Tipo: Polígono Escala: 1: 2.000
7. Shapefile de Cartográfica base de la ciudad de Bogotá.	IDECA	Shapefile(.shp)	Año de Actualización: 2015-2017. Tipo: Polígono Escala: 1: 2.000
<b>Software</b>			
ArcGIS. Extensión: ArcMap 10.3 o versiones superiores			

Anexo 6. Análisis multicriterio para la instalación de estaciones meteorológicas: red de apoyo y monitoreo.

Para la ubicación de las estaciones meteorológicas se recomienda realizar un análisis multicriterio que con base en el uso de sistemas de información geográfica, así como un conjunto de criterios estratégicamente pensados para la ubicación de estaciones meteorológicas, en conjunto forman una herramienta de análisis multicriterio que sirve como base para el análisis espacial y cartográfico, así como la discriminación de las zonas óptimas para la instalación de dichas estaciones meteorológicas; en ciudades como Bogotá, por su importancia regional, es importante contar con una red de tecnología de punta que proporcione datos confiables, que puedan ser usados

como base en la elaboración de proyectos y trabajos de investigación de todo tipo, así como para efectos prácticos de este proyecto: monitorear las zonas propensas a inundaciones en la ciudad.

Las estaciones meteorológicas son lugares que utilizan los instrumentos adecuados para hacer mediciones, observaciones, análisis y pronósticos de los fenómenos meteorológicos en determinado tiempo y lugar. Existen diferentes tipos de estaciones meteorológicas que cumplen funciones específicas como las estaciones pluviométricas, que miden la cantidad de lluvia caída con un pluviómetro, las estaciones climatológicas principales, en la que se realizan observaciones del comportamiento atmosférico en tiempo real de la visibilidad, precipitación, aire, humedad, temperatura del aire y otros fenómenos; estaciones sinópticas, que sirven para medir diversos fenómenos a horas concordantes internacionalmente, y otras.

Para el emplazamiento de cualquier estación meteorológica, la Organización Meteorológica Mundial, ha establecido unos lineamientos, en la Guía de prácticas climatológicas (2011), que indican que las estaciones meteorológicas deben estar ubicadas en sitios planos cubiertos de césped corto, lejos de árboles, edificios, paredes, altas pendientes y hondonadas. Esto permite una correcta disposición y exposición de los instrumentos y una parcela de 9 metros por 6 metros será suficiente para localizar estaciones que midan temperatura y humedad. Si se requiere instalar más instrumentos de medición se necesitaría una parcela de 10 metros por 7 metros.

A pesar de que en Bogotá existen 66 estaciones meteorológicas, que se concentran en la parte central de la ciudad, hay áreas que son aptas para instalar estaciones debido a que tienen las características necesarias para ubicarlas correctamente. Es importante que en la identificación de la ubicación actual de las estaciones se evalúen afectaciones en las mediciones de los instrumentos y los factores que influyen sobre esta problemática con la finalidad de mejorar la precisión de los instrumentos y la calidad de los datos.

Lo anterior quiere decir que los pronósticos del tiempo se basan en mediciones precisas y por lo tanto el presente proyecto tiene el objetivo de ubicar las mejores zonas para establecer estaciones meteorológicas en la ciudad de Bogotá, mediante la búsqueda de información teórica y de capas cartográficas que permitan la creación de criterios, representados en mapas construidos con la

ayuda de un Sistema de Información Geográfica, a continuación se mostrara la descripción de los insumos o base cartográfica mínima para el análisis:

Componente	Capa (Layer)	Modelo	Tipo	Extensión	Escala	Fuente
Antropogénico	Estaciones Meteorológicas Pre-existentes	Vector	Punto	Shapefile		IDEAM
	Red Vial	Vector	Polilinea	Shapefile	1:25.000	IDECA
	Cuerpos de Agua	Vector	Polígono	Shapefile	1:2.000	IDECA
	Acuíferos	Vector	Polígono	Shapefile	1:25.000	SIAT
	Coberturas De la Tierra	Raster	Raster	Raster Dataset	1:25.000	SIAT
Hidrológico y Meteorológico	Temperatura	Raster	Raster	Raster Dataset	1:25.000	IDECA
	Precipitación	Raster	Raster	Raster Dataset	1:25.000	IDECA
	Elevación	Raster	Raster	Raster Dataset	1:25.000	USGS

*Determinación de criterios Antropogénicos:*

El objetivo de implementar este conjunto de criterios es principalmente que las estaciones se logren ubicar en donde las condiciones tanto: viabilidad de implementación, operatividad, movilidad, seguridad entre otras, sean las más adecuadas tanto para el personal técnico y de mantenimiento, como para la recolección de datos por parte de los equipos tecnológicos especializados; lejos de zonas bajo amenaza de inundación ( problema muy frecuente en la ciudad de Bogotá, debido a la cantidad de humedales y acuíferos, así como por la deficiencia en el alcantarillado), así como en

zonas con coberturas vegetales prominentes o espesas, que resulten ser desfavorables para la medición de los parámetros Hidrológicos y meteorológicos, los criterios se listarán a continuación:

1. **Cercanía a vías principales:** las estaciones meteorológicas deben estar a no más de 500 m de vías de acceso principales con el objetivo de garantizar el fácil acceso por parte del personal técnico y de mantenimiento.
2. **Cuerpos de Agua:** las estaciones meteorológicas deben estar ubicadas a más de 500 m de los principales cuerpos de agua, con el fin de evitar riesgo de daños en equipos, en caso de una eventual inundación.
3. **Lejanía a estaciones:** preexistentes: las estaciones meteorológicas deben estar a más de 1.500 m de las estaciones existentes que estén vigentes.
4. **Coberturas del suelo:** las estaciones meteorológicas no deben estar ubicadas dentro de coberturas de la tierra que pueda afectar de manera indirecta las mediciones por parte de los equipos, por lo tanto es recomendable ubicarlas dentro de coberturas de pastizales abiertos y demás coberturas de este tipo.

#### ***Determinación de criterios hidrológicos y meteorológicos:***

Estos criterios se establecen con el objetivo de que las estaciones se ubiquen en zonas en donde factores como: la precipitación, temperatura y humedad, sean representativos con el fin de obtener de la misma forma datos representativos condicionantes en la generación de por ejemplo: islas de calor o microclimas urbanos.

1. **Precipitación y Temperatura:** las estaciones meteorológicas deben ubicarse en lugares en donde los valores para dichas variables sean representativos.
2. **Pendiente:** Las estaciones meteorológicas deben ubicarse dentro de una pendiente no mayor a 7% (según los rangos de clasificación del IGAC), con el objetivo, de que las

estaciones sean capaces de recibir la incidencia directa de la radiación solar, y que no se afecten las mediciones por interferencias de algún tipo.

***Obtención de pendientes del Bogotá D.C mediante la utilización de herramientas de análisis espacial:***

***Slope.***

Con la Imagen satelital obtenida del USGS, se puede generar un mapa de pendientes usando la herramienta Slope que se encuentra en ArcToolbox > Spatial Analyst Tools > Surface > Slope. Dentro de esta herramienta se tuvo en cuenta la opción de Percent Rise, para que los valores de las pendientes fueran medidos en porcentajes, además se consideró cambiar algunos ambientes de trabajo, con el uso de la herramienta Environments, como el sistema de coordenadas, MAGNA Colombia Bogotá, el tamaño de la celda, 30x30m, y la máscara de geoprocesamiento basada en el insumo de los límites municipales.

Al tener listo el mapa de pendientes con el área de estudio establecida, se debe hacer una reclasificación de los valores de las pendientes con la herramienta Reclassify, encontrada en ArcMap dentro de ArcToolbox > Spatial Analyst Tools > Reclass > Reclassify. En esta parte se editaron manualmente las clases, mediante la opción de Classify, según las 8 clases establecidas por el IGAC:

Rangos Para Pendiente del IGAC.

<b>CLASE</b>	<b>RANGO DE PENDIENTES (%)</b>
<b>1</b>	0-3
<b>2</b>	3-7
<b>3</b>	7-12
<b>4</b>	12-25
<b>5</b>	25-50

<b>6</b>	50-100
<b>7</b>	100 - último valor
<b>8</b>	Otros valores

En este proceso también se tuvo en cuenta los ambientes de trabajo como el sistema de coordenadas, MAGNA Colombia Bogotá, el tamaño de la celda, 30x 30m y la máscara de geoprocesamiento.

Por otro lado se convirtieron varias de las capas entendidas como cartografía base a formato raster a través de la herramienta Polygon to Raster, la cual está en ArcToolbox > Conversion Tools > To Raster > Polygon to Raster. En esta opción el valor del campo que se utilizó fue la de los nombres de las coberturas, igualmente, se cambiaron los ambientes de trabajo con los mismos que se deben tener en cuenta para el mapa de pendientes.

Posteriormente con el mapa de pendientes y el mapa de coberturas, con las mismas características de formato, sistema de coordenadas, tamaño de celda y máscara de geoprocesamiento, se utilizó la herramienta de modelo de superposición por pesos ponderados, para crear la superficie de fricción, que sirvió como base para el análisis.

En la herramienta Weighted Overlay, del ArcToolbox > Spatial Analyst Tools > Overlay, se subieron los 2 mapas resultantes obtenidos de los procesos descritos anteriormente, se estimó una escala de evaluación de 1 a 4 por 1, donde 1 es el menor coste y valor permitido, 2 es el valor de menor coste y una restricción más alta que el valor 1, 3 es un coste y restricción más alta y 4 es un mayor coste y restricción más alta para la determinación de las zonas óptimas para la instalación de una estación meteorológica. Los pesos que se usaron para el modelo fueron 60-40, respectivamente.

Para este caso se utilizó el siguiente rango de normalización de los datos: 1 to 4 by 1 siendo 1 el valor de menor coste o valor permitido, 2 como el segundo valor de menor coste, con una restricción más alta que para el primer valor, 3 como un valor con un coste y restricción mayor, y 4 como el valor de mayor coste y restricción para la determinación de las zonas óptimas para la

instalación de una estación meteorológica de cualquier tipo; cabe destacar que también se le puede asignar valores a los campos como restringido, para casos específicos, si lo que se desea es que definitivamente dicho campo o factor sea totalmente excluido como una zona optima que cumpla con las condiciones necesarias.

Los pesos ponderados para el mapa de pendientes y coberturas de la tierra son:

PENDIENTES	
Clases <sup>3</sup>	Rango de Normalización
1	1
2	1
3	RESTRICTED
4	RESTRICTED
6	RESTRICTED
7	RESTRICTED
8	RESTRICTED
COBERTURAS	
Mosaico de pastos con espacios naturales	2
Pastos Limpios	1
Mosaico de pastos y cultivos	3
Tierras desnudas y degradadas	1
Herbazal	2
Mosaico de Cultivos, pastos y espacios naturales	3
Arbustal	1
Cuerpos de Agua Artificiales	RESTRICTED
Zonas industriales o comerciales	RESTRICTED
Plantación Forestal	RESTRICTED

Mosaico de Cultivos con espacios naturales	2
Bosque fragmentado	4
Bosque Abierto	4
Vegetación secundaria o en transición	4
Bosque denso	4
Bosque de galería o ripario	RESTRICTED
Pastos enmalezados	2
Pastos arbolados	4
Cultivos permanentes arbustivos	4
Cultivos permanentes herbáceos	2
Tejido urbano continuo	1
Tejido urbano discontinuo	1
Mosaico de cultivos	4
Ríos (50m)	RESTRICTED
Lagunas, lagos y ciénagas naturales	RESTRICTED
Zonas de extracción minera	RESTRICTED
Otros cultivos transitorios	4
Aeropuertos	RESTRICTED

## REFERENCIAS

- Aisha, T. S., Wok, S., Manaf, A. M. A., & Ismail, R. (2015). Exploring the Use of Social Media During the 2014 Flood in Malaysia. *Procedia - Social and Behavioral Sciences*, 211(September), 931–937. <https://doi.org/10.1016/j.sbspro.2015.11.123> Congreso de la Republica. (24 de Abril de 2012). *Poder Público - Rama Legislativa*. Obtenido de LEY 1523 DE 2012: [https://www.arlsura.com/files/ley1523\\_2012.pdf](https://www.arlsura.com/files/ley1523_2012.pdf)
- Aguilar, L. J. (2013). *Big Data: Análisis de grandes volúmenes de datos en organizaciones*. Mexico D.F: Alfaomega.
- Alcaldía Mayor de Bogotá. (27 de Septiembre de 2010). *Regimen legal de Bogotá*. Obtenido de DECRETO 413 DE 2010: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=40442>
- Aler, R. (2009). *Tutorial Weka 3.6.0*. Obtenido de <http://ocw.uc3m.es/ingenieria-informatica/herramientas-de-la-inteligencia-artificial/contenidos/transparencias/TutorialWeka.pdf>
- Alvear, J. O. (Noviembre de 2018). *Arboles de decision y Random Forest*. Obtenido de <https://bookdown.org/content/2031/ensambladores-random-forest-parte-i.html#ventajas-de-random-forest>
- Arango, C., Dorado, J., D., G., & Ruiz, J. F. (2012). *IDEAM*. Obtenido de CAMBIO CLIMÁTICO MÁS PROBABLE PARA COLOMBIA A LO LARGO DEL SIGLO XXI RESPECTO AL CLIMA PRESENTE: <http://www.ideam.gov.co/documents/21021/21138/Escenarios+Cambio+Climatico+%28Ruiz%2C+Guzman%2C+Arango+y+Dorado%29.pdf/fe5d64fb-3a82-4909-a861-7b783d0691cb>
- ArcgisPro. (Abril de 2019). *Poligonos de Thiessen*. Obtenido de Crear polígonos Thiessen: <https://pro.arcgis.com/en/pro-app/tool-reference/analysis/create-thiessen-polygons.htm>
- Ballis, T. (s.f.). *Tipos y fases de los desastres*. Obtenido de <http://www.ispch.cl/sites/default/files/1.-%20Tipos%20y%20Fases%20de%20Desastres.pdf>
- Banco Mundial Colombia. (2012). *Análisis de la gestión del riesgo de desastres en Colombia: Un aporte para construcción de políticas públicas*. Obtenido de [www.gestiondesriesgo.gov.co/sigpad/archivos/GESTIONDELRIESGOWEB.pdf](http://www.gestiondesriesgo.gov.co/sigpad/archivos/GESTIONDELRIESGOWEB.pdf)
- Beltrán, J. P. (14 de Febrero de 2019). *Bogotá región: crecimiento urbano en la consolidación del territorio metropolitano*. Obtenido de Universidad Francisco José de Caldas: [http://www.institutodeestudiosurbanos.info/dmdocuments/cendocieu/coleccion\\_digital/Practicas\\_Ancestrales\\_Sabana/Bogota\\_Region\\_Crecimiento-Preciado\\_J.pdf](http://www.institutodeestudiosurbanos.info/dmdocuments/cendocieu/coleccion_digital/Practicas_Ancestrales_Sabana/Bogota_Region_Crecimiento-Preciado_J.pdf)

- Bertoglio, O. J. (1996). *Introducción a la teoría general de sistemas*. México: Limusa.
- Bertoni, J. C. (2013). *La problemática de las inundaciones urbanas: el caso de la cuenca Matanza-Riachuelo*. Córdoba, Argentina: Programa Hidrológico Internacional PHI/LAC.
- Bogotá, A. M. (20 de Febrero de 2019). *Línea de Atención 123*. Obtenido de NUSE: <https://scj.gov.co/landing/linea-123/>
- CAR. (Febrero de 2014). *Corporación Autónoma Regional*. Obtenido de ADECUACIÓN HIDRÁULICA Y RECUPERACIÓN AMBIENTAL DEL RÍO BOGOTÁ: <https://www.car.gov.co/uploads/files/5aec3381e652.pdf>
- CENAPRED. (Agosto de 2009). Inundaciones. *CENAPRED*, 3-38. Obtenido de centro nacional de prevención de desastres: [http://www.proteccioncivil.gob.mx/work/models/ProteccionCivil/Resource/377/1/images/folleto\\_i.pdf](http://www.proteccioncivil.gob.mx/work/models/ProteccionCivil/Resource/377/1/images/folleto_i.pdf)
- CENAPRED. (2019). *Inundaciones*. Obtenido de <https://www.cenapred.gob.mx/es/Publicaciones/archivos/3-FASCCULOINUNDACIONES.PDF>
- CEPAL. (2007). *Información para la gestión de riesgo de desastres: Estudio de caso de cinco países: Colombia*. Obtenido de <https://repositorio.cepal.org/handle/11362/25930>
- CEPAL. (Enero de 2012). *Valoración de daños y pérdidas. Ola invernal en Colombia 2010-2011*. Obtenido de [https://repositorio.cepal.org/bitstream/handle/11362/37958/1/OlainvernalColombiaBIDCEPAL\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/37958/1/OlainvernalColombiaBIDCEPAL_es.pdf)
- CLGR-CC. (Junio de 2018). *caracterización general de escenarios de riesgo*. Obtenido de <https://www.idiger.gov.co/documents/220605/314085/Identificaci%C3%B3n+y+priorizaci%C3%B3n.pdf/24386a78-ea2b-4abe-9516-9b9c37955fc4>.
- Concejo de Bogotá. (2012). *ACUERDO 489*. Obtenido de Plan de Desarrollo Bogotá Humana: <http://www.saludcapital.gov.co/CTDLab/Antecedentes%20Normativos/Acuerdo%20489-12%20Plan.pdf>
- Concejo de Bogotá. (30 de Septiembre de 2014). *ACUERDO 001 DE 2014*. Obtenido de Alcaldía Mayor de Bogotá: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=59695>
- Concejo de Bogotá D.C. (7 de Diciembre de 2013). *Acuerdo 546 de 2013*. Obtenido de Alcaldía Mayor de Bogotá: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=56152>

- Congreso de Colombia. (2012). *Ley 1523 de 2012*. Obtenido de <http://www.ideam.gov.co/documents/24189/390483/11.+LEY+1523+DE+2012.pdf/4e93527d-3bb8-4b53-b678-fbde8107d340?version=1.2>
- Corso, C. L. (s.f.). *Aplicación de algoritmos de clasificación supervisada usando Weka*. Obtenido de [https://www.academia.edu/9538896/Aplicaci%C3%B3n\\_de\\_algoritmos\\_de\\_clasificaci%C3%B3n\\_supervisada\\_usando\\_Weka](https://www.academia.edu/9538896/Aplicaci%C3%B3n_de_algoritmos_de_clasificaci%C3%B3n_supervisada_usando_Weka)
- DANE. (2011). *Registro de damnificados por la emergencia invernal 2010 - 2011*. Obtenido de <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/registros-emergencia-invernal>
- DANE. (2018). *Boletín Preliminar*. Obtenido de Herramientas técnicas: Censo Nacional 2018: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivenda-2018/informacion-tecnica>
- Decreto 141 de 2011 Nivel Nacional, Sentencia C-276 de 2011 (Corte Constitucional).
- D'epenoux, É., KOGUC, S., & Hennebert, J. (22 de Enero de 2018). *Las Inundaciones*. afpes.
- DIAN. (2015). *Estadísticas*. Obtenido de Cifras Nacionales: Crecimiento y desarrollo: <https://www.dian.gov.co/dian/cifras/Paginas/estadisticas.aspx>
- Díaz-Bravo, L., Torruco-García, U., Martínez-Hernández, M., & Varela-Ruiz, M. (2013). *La entrevista, recurso flexible y dinámico*. Obtenido de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-50572013000300009](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572013000300009)
- EIRD y OCAH. (2008). *Preparación ante los desastres para una respuesta eficaz. Conjunto de directrices e indicadores para la aplicación de la prioridad 5 del Marco de Hyogo*. Obtenido de [https://www.unisdr.org/files/2909\\_OCHADisasterpreparednesseffectiveresponseSPA.pdf](https://www.unisdr.org/files/2909_OCHADisasterpreparednesseffectiveresponseSPA.pdf)
- Galindo, W. G. (2013). *dinámica de las construcciones por usos de la localidad de Kennedy en los años 2002 y 2012*. Obtenido de <https://www.catastro bogota.gov.co/sites/default/files/10.pdf>
- Idarraga, F. L. (Julio de 2010). *Respuestas y propuestas ante el riesgo de inundación de las ciudades colombianas*. Obtenido de <http://www.redalyc.org/html/1210/121015012007/>
- IDC. (2012). *Resumen ejecutivo - Big Data: un mercado emergente*. Obtenido de <https://www.diarioabierto.es/wp-content/uploads/2012/06/Resumen-Ejecutivo-IDC-Big-Data.pdf>
- IDEAM. (2012). *ANÁLISIS DEL IMPACTO DEL FENÓMENO “LA NIÑA” 2010-2011 EN LA HIDROCLIMATOLOGÍA DEL PAÍS*. Bogota D.C.

- IDEAM. (30 de Enero de 2019). *Laboratorio de Calidad Ambiental*. Obtenido de FEWS COLOMBIA: <http://www.ideam.gov.co/web/agua/fews>
- IDEAM. (s.f.). *Amenazas inundación*. Obtenido de <http://www.ideam.gov.co/web/agua/amenazas-inundacion>
- IDEAM, & MINAMBIENTE. (2014). *Estudio Nacional del agua*. Bogotá.
- IDIGER. (01 de Agosto de 2016). *Instituto Distrital de Gestion de Riesgos y Cambio Climatico*. Obtenido de HISTORIA: <http://www.idiger.gov.co/historia>
- IDIGER. (2017). *Sistema Distrital de Gestión de Riesgos y Cambio Climatico*. Obtenido de Marco de Actuación: Estrategia Distrital para la Respuesta a Emergencias: <https://www.sire.gov.co/guias>
- IDIGER. (25 de Junio de 2018). *INSTITUTO DISTRITAL DE GESTIÓN DE RIESGOS Y CAMBIO CLIMÁTICO*. Obtenido de PROYECTO ACTUALIZACIÓN COMPONENTE DE GESTIÓN DEL RIESGO PARA LA REVISIÓN ORDINARIA Y ACTUALIZACIÓN DEL PLAN DE ORDENAMIENTO TERRITORIAL.: <http://www.sdp.gov.co/sites/default/files/4-DOCUMENTO-TECNICO-DE-SOPORTE/Gestion%20del%20Riesgo.%20Amenazas%20inundacion%20Urbano.pdf>
- IDIGER. (27 de Noviembre de 2018). *Instituto Distrital de Gestion de Riesgos y Cambio Climatico*. Obtenido de Caracterización General de Escenario de Riesgo de Inundación por Desbordamiento: <http://www.idiger.gov.co/rinundacion>
- IDIGER. (21 de diciembre de 2018). *Organigrama*. Obtenido de <http://www.idiger.gov.co/organigrama#EMER>
- IDIGER. (2019). *Caracterización general del escenario de riesgo por inundación*. Obtenido de <https://www.idiger.gov.co/rinundacion>
- IGAC, IDEAM, & DANE. (2011). *REPORTE FINAL DE ÁREAS AFECTADAS POR INUNDACIONES 2010 – 2011*. Bogota.
- INDECI. (20 de Enero de 2019). *Institucion nacional de defensa civil*. Obtenido de Diferencia entre Danmificado y afectado : <http://eltiempo.pe/wp-content/uploads/2017/03/indeci.png>
- ISO. (2009). *ISO Guide 73:2009 Preview*. Obtenido de <https://www.iso.org/obp/ui/#iso:std:iso:guide:73:ed-1:v1:en>
- López-Roldán, P., & Fachelli, S. (2016). *METODOLOGÍA DE LA INVESTIGACIÓN SOCIAL CUANTITATIVA*. Obtenido de [https://ddd.uab.cat/pub/caplli/2016/163567/metinvsoccua\\_a2016\\_cap2-3.pdf](https://ddd.uab.cat/pub/caplli/2016/163567/metinvsoccua_a2016_cap2-3.pdf)

- Losada, H. F., Rubiano, M. F., & Muñoz, J. C. (Diciembre de 2016). *Universidad Distrital Francisco José de Caldas*. Obtenido de Revista Científica: <https://revistas.udistrital.edu.co/ojs/index.php/revcie/article/download/11087/11929/>
- Lozada, A. M. (7 de Mayo de 2017). *El Tiempo*. Obtenido de Línea 123 lleva diez años trabajando con la misma tecnología: <https://www.eltiempo.com/bogota/linea-123-lleva-diez-anos-con-la-misma-tecnologia-85568>
- Marqués, M. P. (2015). *Big Data. Técnicas, herramientas y aplicaciones*. México D.F: Alfaomega.
- Mena, L. G. (11 de Abril de 2016). *Las TIC como herramienta transversal*. Obtenido de <http://docentes.unibe.edu.do/las-tic-como-herramienta-transversal/>
- Merino, R. F., & Chacón, C. I. (23 de Septiembre de 2017). *Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python*. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/6230447.pdf>
- Merino, R. F., & Chacón, C. I. (23 de Septiembre de 2017). *Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python*. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/6230447.pdf>
- Minambiente. (15 de diciembre de 2010). *Ministerio de Ambiente, vivienda y Desarrollo Sostenible*. Obtenido de <http://www.minambiente.gov.co//contenido/contenido.aspx?conID=6714&catID=1155>
- Ministerio de cultura. (2014). *Cartilla básica de gestión del riesgo para patrimonio material e infraestructura cultural*. Obtenido de <http://www.mincultura.gov.co/sitios/todos-por-el-choco/cartilla.pdf>
- PICÓN, R. J., GUZMÁN, I. D., & ROLDÁN, J. B. (2012). *Memoria técnica: Evaluación, análisis y seguimiento a las afectaciones por inundaciones asociadas al fenómeno de la niña 2010-2011*. Bogotá.
- PNUD. (2014). *Reducción de Riesgos en Tabasco: Gobernanza del Riesgo y Resiliencia Territorial*. Tabasco, México: PUBLICACIONES DE MEDIO AMBIENTE Y DESARROLLO SUSTENTABLE.
- Ramirez, A. (22 de Enero de 2018). *InformaBTL*. Obtenido de ¿Cuáles son las diferencias entre consumidor, cliente, comprador y usuario?: <https://www.informabtl.com/cuales-las-diferencias-consumidor-cliente-comprador-usuario/>
- Revista semana. (28 de Abril de 2015). *Río Bogotá, contaminación sin tregua*. Obtenido de Medio Ambiente: <https://sostenibilidad.semana.com/medio-ambiente/articulo/rio-bogota-contaminacion-tregua/32929>
- Rodríguez, E., Corrales, A. L., & Rodríguez, R. (2016). *AVANCES DE INVESTIGACIÓN EN LA MEJORA DE LA EDUCACIÓN EN VALORES Y FORMACIÓN DOCENTE*

*INTEGRAL*. Obtenido de [https://www.academia.edu/34635530/La\\_Comunicaci%C3%B3n\\_Mediada\\_por\\_Computadora.\\_Origen\\_definici%C3%B3n\\_caracter%C3%ADsticas\\_e\\_investigaciones\\_educativas](https://www.academia.edu/34635530/La_Comunicaci%C3%B3n_Mediada_por_Computadora._Origen_definici%C3%B3n_caracter%C3%ADsticas_e_investigaciones_educativas)

- Romero, S. (2017). *El 'big data', aliado para afrontar catástrofes naturales como el ciclón australiano 'Debbie'*. Obtenido de BBVA: <https://www.bbva.com/es/big-data-aliado-afrontar-catastrofes-naturales-ciclon-australiano-debbie/>
- Sande, J. C. (Junio de 2018). *Análisis de Sentimientos en Twitter*. Obtenido de <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/81435/6/jsobrinosTFM0618memoria.pdf>
- SAV y DGPC. (2015). *Conceptos básicos de gestión de riesgos*. Obtenido de <http://dipecholac.net/docs/herramientas-proyecto-dipecho/el-salvador/C1-CONCEPTOS-BASICOS-DE-GRD.pdf>
- Secretaria Distrital de Planeación. (2012-2016). *Plan de Desarrollo "Bogota Humana"*. Obtenido de Alcaldía Mayor de Bogotá: <http://www.sdp.gov.co/gestion-a-la-inversion/planes-de-desarrollo-y-fortalecimiento-local/planes-de-desarrollo-local/bogota-humana>
- Sentencia C-216 de 2011 Corte Constitucional, expediente RE-197 (29 de Marzo de 2011).
- Sheriff's Office Jefferson County. (s.f.). *Equipo de Gestión de Incidentes*. Obtenido de <https://www.jeffco.us/490/Incident-Management-Team>
- SNGRD. (2017). *Plan sectorial de contingencias frente a de lluvias 2016 - 2018 y posible fenómeno de La Niña*. Obtenido de [http://www.minambiente.gov.co/images/OrdenamientoAmbientalyTerritorialyCoordinaciondelSIN/pdf/boletines/Plan\\_de\\_contingencia\\_temporada\\_lluvias-2017.pdf](http://www.minambiente.gov.co/images/OrdenamientoAmbientalyTerritorialyCoordinaciondelSIN/pdf/boletines/Plan_de_contingencia_temporada_lluvias-2017.pdf)
- UNESCO/OMM. (1974). *UNESDOC*. Obtenido de UNESCO Digital Library: <https://unesdoc.unesco.org/ark:/48223/pf0000221862>
- UNGRD. (2017). *Terminología sobre gestión del riesgo de desastres y fenómenos amenazantes*. Obtenido de <https://repositorio.gestiondelriesgo.gov.co/bitstream/handle/20.500.11762/20761/Terminologia-GRD-2017.pdf;jsessionid=F84E67CBD61B411BAC320F83DC39DEEE?sequence=2>
- UNGRD. (2018). *Guía para la conformación de la estructura organizacional de gestión del riesgo de desastres en las entidades territoriales*. Obtenido de <http://online.pubhtml5.com/pxou/jpxi/#p=2>
- UNGRD. (2018). *Impactos de los eventos recurrentes y sus causas en Colombia*. Obtenido de <http://online.pubhtml5.com/pxou/aqzy/#p=1>

- UNGRD. (Junio de 2018). *Manual de comunicación en emergencia*. Obtenido de <http://online.pubhtml5.com/pxou/btsv/#p=14>
- UNI. (2019). *Machine Learning*. Obtenido de <http://mim.promexico.gob.mx/work/models/mim/templates-new/Publicaciones/Notas/Machine-Learning.pdf>
- Unidad Nacional para la Gestión del Riesgo de Desastres. (2016). *Plan Nacional de Gestión de Riesgo de Desastres. Una estrategia de desarrollo 2015-2025*. Obtenido de <http://portal.gestiondelriesgo.gov.co/Paginas/Plan-Nacional-Gestion-Riesgo-de-Desastres.aspx>
- UNISDR. (2009). *Terminología sobre reducción del riesgo de desastres*. Obtenido de [https://www.unisdr.org/files/7817\\_UNISDRTerminologySpanish.pdf](https://www.unisdr.org/files/7817_UNISDRTerminologySpanish.pdf)
- Universidad de Murcia. (s.f.). *Mi primera hora con Eclipse*. Obtenido de <https://www.um.es/docencia/barzana/DAWEB/Mi-primera-hora-con-Eclipse.pdf>
- Universidad de Salamanca España. (20 de Febrero de 2019). *Hidrología Usal*. Obtenido de [http://hidrologia.usal.es/practicas/Pneta\\_SCS/Pneta\\_SCS\\_fundam.pdf](http://hidrologia.usal.es/practicas/Pneta_SCS/Pneta_SCS_fundam.pdf)
- UPM. (11 de Octubre de 2017). *¿Cómo ayuda el "Big Data" en desastres naturales?* Obtenido de <https://www.upm.es/e-politecnica/?p=8946>
- Velandia, J. C. (2014). *LA GESTIÓN DEL RIESGO DE DESASTRES EN LAS INUNDACIONES DE COLOMBIA: UNA MIRADA CRÍTICA*. Obtenido de [https://repository.ucatolica.edu.co/bitstream/10983/2128/1/Gesti%C3%B3n\\_riesgo\\_desastres\\_inundaciones\\_%20Colombia\\_mirada-cr%C3%ADtica.pdf](https://repository.ucatolica.edu.co/bitstream/10983/2128/1/Gesti%C3%B3n_riesgo_desastres_inundaciones_%20Colombia_mirada-cr%C3%ADtica.pdf)