

## Big Data y áreas de oportunidad para la proyección del Sistema Inteligente de Transporte en Bogotá, Colombia

## Big Data and Areas of Opportunity for the Projection of the Intelligent Transportation System in Bogotá, Colombia

Nancy Edith Ochoa Guevara<sup>1</sup>

[nancy.ochoa@uniremington.edu.co](mailto:nancy.ochoa@uniremington.edu.co)

<https://orcid.org/0000-0002-4533-4990>

Sandra Patricia Ochoa Guevara<sup>2</sup>

[sandra.ochoa@mail.udes.edu.co](mailto:sandra.ochoa@mail.udes.edu.co)

<https://orcid.org/0000-0001-6089-1229>

Pedro Adolfo Garzón Martínez<sup>3</sup>

[67172513@unitec.edu.co](mailto:67172513@unitec.edu.co)

<https://orcid.org/0000-0002-4753-4558>

<https://doi.org/10.22209/rhs.v11n1a09>

**Recibido:** enero 25 de 2023.

**Aceptado:** mayo 8 de 2023.

**Para citar:** Ochoa Guevara, N. E., Ochoa Guevara, S. P., & Garzón Martínez, P. A. (2023). Big Data y áreas de oportunidad para la proyección del Sistema Inteligente de Transporte en Bogotá, Colombia. *RHS-Revista Humanismo y Sociedad*, 11(1), 1-17.

<https://doi.org/10.22209/rhs.v11n1a07>

### Resumen

Hoy en día, en las grandes ciudades de Colombia, en especial en Bogotá, y debido al crecimiento de su población (9,3 millones con la llegada de inmigrantes), se exige una demanda de aporte a la proyección de sistemas inteligentes de transporte públicos y privados como un logro de la política de movilidad de la administración de la Bogotá Humana. De ahí surge el interrogante: ¿cuál es el desafío y las áreas de oportunidad de adaptar un Big Data en la proyección de un Sistema Inteligente de Transporte para todos los ciudadanos en Bogotá? A partir de esta pregunta, se propone determinar los aportes que el Big Data ofrece como centro de acopio en la proyección de un sistema inteligente para la ciudad. La indagación se plantea desde un enfoque cualitativo y un estudio descriptivo. Se incluye la revisión de algunos estudios realizados mediante las técnicas del Big Data y del análisis de datos de contenido de la estructura organizada de estos por la Secretaría Distrital de Movilidad en Bogotá. Los resultados permiten orientar los aportes del Big Data después de analizar la estructura de indicadores que ofrecen estos el conjunto de datos. A partir de estos, se encuentran brechas y vacíos preocupantes para el Sistema Inteligente de Transporte que se espera en el futuro para Bogotá.

1 Corporación Universitaria Remington, Medellín, Colombia.

2 Universidad de Santander (UDES), Bucaramanga, Colombia.

3 Corporación universitaria Unitec, Bogotá, Colombia.

**Palabras clave:** analítica de datos, brechas de datos, indicadores, tecnología, volumen de datos, metodología.

---

## Abstract

Today, the large cities of Colombia – especially Bogotá, due to the growth of its population (9.3 million with the arrival of immigrants) – demand the projection of intelligent public and private transport systems, as an achievement of the mobility policy of the Bogota Humana administration. Hence, this question arises: What are the challenges and areas of opportunity of adapting Big Data to project an Intelligent Transportation System for all citizens in Bogotá? Based on this question, our aim is to determine the contributions that Big Data offers as a collection center for the projection of an intelligent system for the city. Our research was proposed with a qualitative approach and a descriptive study. The review of some studies developed using Big Data techniques and content data analysis of their organized structure by the District Mobility Secretariat in Bogotá was included. The results allow guiding the contributions of Big Data after analyzing the structure of indicators offered by the data set. From these, we found gaps and voids that are concerning for the Intelligent Transportation System that is expected in the future for Bogotá.

**Keywords:** Data analytics, Data gaps, Indicators, Technology, Data volume, Methodology.

## Introducción

Los *Intelligent Transport Systems* [ITS] tienen como propósito el control del tránsito y el transporte; revisan, analizan, evalúan y proponen soluciones para la congestión vehicular, la accidentalidad y la contaminación ambiental, entre otros. Estos sistemas le apuntan a una gestión eficiente, sostenible y sustentable de la infraestructura vial, en especial en las grandes ciudades.

Según Molina y Espitia (2016), este objetivo se logra con el apoyo de herramientas de última generación enmarcadas en la industria 4.0 y la aplicación de la inteligencia artificial con datos estructurados y no estructurados, extraídos desde los dispositivos y procesados en esquemas de computación en la nube, internet de las cosas (IoT) y algoritmos de aprendizaje desde el Big Data. Esto representa un gran aporte de las tecnologías de la información y la comunicación [TIC] para las grandes ciudades del mundo (Zanoon *et al.*, 2017).

Por tanto, el Big Data, como mencionan Arango *et al.* (2015), se aprovecha en gran parte de los dispositivos y periféricos electrónicos que se encuentran dentro y fuera de la ciudad de Bogotá. Gracias a estas conexiones, detecta datos que aportan a las predicciones en tiempo real, al ser procesados y analizados para identificar y visualizar una situación específica.

Según Cacho-Elizondo y Lázaro (2018), la transformación digital “se trata del proceso de modificar a una organización usando

herramientas innovadoras, así como adoptar tecnología de punta y, al mismo tiempo, cambiar radicalmente la cultura corporativa con el propósito de adoptar nuevos modelos de operación y de negocio” (Rincón-García, p. 33).

De ahí que, la empresa Moovit (2022) revela las últimas tendencias de transporte a partir de las búsquedas realizadas por los usuarios. Allí Rincón-García *et al.* (2020), ofrece una radiografía de cómo se mueven las personas en las ciudades grandes del mundo. En este caso se filtra la ciudad de Bogotá en Colombia y se enuncian los siguientes resultados con una estadística a enero del 2022.

El promedio de tiempo que dedica la gente a desplazarse en transporte público es de 67 minutos. El 64 % de estos usuarios se toman un promedio superior de 2 horas al día en sus desplazamientos. La media de tiempo que una persona espera en una estación de transporte público es de 21 minutos, mientras que el 56 % excede este lapso. Sin embargo, el 47 % de los usuarios de transporte en la capital viajan más de 12 km en una sola dirección y la distancia media que una persona camina desde su última parada a su destino es de 1,09 Km, y el 43 % camina más de 1 Km para llegar a su destino.

Por lo tanto, para Kitzmann *et al.* (2022) esa información, por sus propiedades, demanda de técnicas de recolección, almacenamiento, gestión e interpretación para que se logren crear modelos con márgenes de error muy bajos. Este conjunto de habilidades entre hombre, datos y máquina se conoce como análisis de datos. Su naturaleza es describir la “Administración”, es decir, la política institucional para poder extraer, unificar y estandarizar la información de distintas fuentes, fortaleciendo la toma de decisiones desde la técnica de

la predicción. Todo esto no sería posible sin intervención humana, recolección de distintas fuentes de datos como pueden ser radiografías, sensores, tablas de datos, videos de plataformas como Facebook, etc. o de cualquier otra fuente; así como la ayuda de tecnologías desarrolladas en programas (*software*) y herramientas (*hardware*) que dan forma, crean y transforman nuevos datos desde su estructura y no estructura bajo la técnica del Big Data (Suat-Rojas, 2021).

De acuerdo con Vásquez (2021), los sistemas de transporte se deben adaptar de acuerdo con la dinámica y globalización de la economía, infraestructura y, sobre todo, con el crecimiento poblacional. La innovación constante e integración de todo tipo de ciencias y tecnologías, como el *Internet of Things* (IoT) (internet de las cosas), darán paso a infraestructuras más robustas, interconectadas e inteligentes donde los datos de sensores, drones, cámaras, vehículos, actores (entidades u organizaciones) y sistemas de posicionamiento Global-GPS brindarán un análisis más sofisticado de la información que permitirá un mejor desempeño de los *Intelligent Transport Systems*.

El análisis y gestión de datos en los sistemas de transporte trae consigo ventajas que suponen mejorías en aspectos cotidianos como el hecho de poder hacer pagos electrónicos al instante en peajes, sin intervención de un operador y sin efectivo; reconocimiento facial en seguridad vial; reducción en la tasa de accidentalidad o en consumo de combustible usando sensores de medición para ver aspectos como velocidad, alteraciones en las rutas, tiempos de salida, llegada a estaciones. Todo esto permite la supervisión, administración y seguimiento sobre los protagonistas viales, lo que optimiza las rutas la productividad,

competitividad de todo el sector, y brinda una mejoría en la experiencia e información hacia el usuario (Cheng, *et al.*, 2015).

Como objetivos en la analítica de datos, la técnica del Big Data y los ITS en el planeta logran caracterizar el desarrollo social que se ve reflejado en la calidad de vida de los usuarios del sistema, la seguridad en las vías, los peatones, la eficiencia y mejora en el transporte público por carreteras, la reducción en problemas de salud de los conductores y peatones.

La importancia de esta investigación radica en dilucidar cuál es el desafío y las áreas de oportunidad de adaptar un Big Data en la proyección de un Sistema Inteligente de Transporte para todos los ciudadanos en Bogotá.

De allí que el objetivo es identificar el desafío y las oportunidades en la proyección de un ITS en Colombia, en especial en la ciudad de Bogotá, por medio de la búsqueda de estudios y proyectos sobre los ITS para las grandes ciudades; de igual forma con el análisis de la estructura de contenido de siete grupos de datos al azar extraídas de la Secretaría Distrital de Movilidad en Bogotá.

Se parte de algunos modelos y métodos utilizados desde la tecnología de última generación en algunos países guía para el ITS de Bogotá, resaltando áreas primordiales que se cubrirán y aportarán de forma exitosa en la proyección del ITS. De igual forma, el análisis de los datos arroja indicadores preocupantes al revisar

su estructura desde el contenido, que resulta vacía y redundante, con falta de integridad y bastantes valores en cero (0).

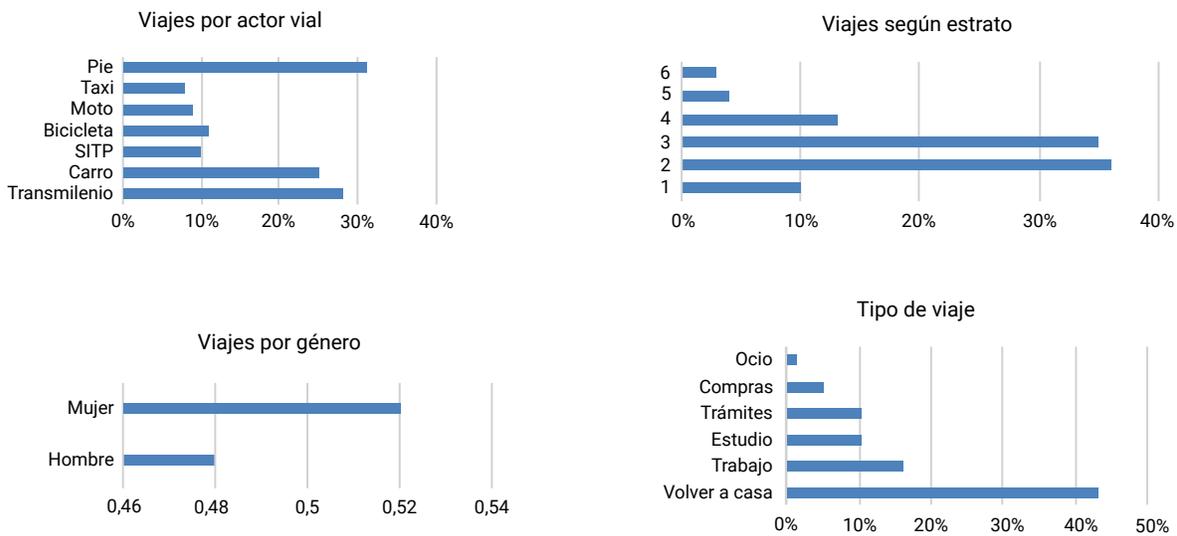
## Marco conceptual

### Impacto en el sector transporte

Los volúmenes incrementales de acuerdo con [Bbvaapimarket.com](http://Bbvaapimarket.com) (2020) en la compra de carros, motos y el uso de medios de transporte como taxis, carros particulares, públicos, tráfico pesado, obras en vía, cierre de carriles y demás, son los factores más notorios en de la administración del tráfico. Por eso, en la adopción de nuevas tecnologías, soluciones y planes de mejora, se encuentra el uso compartido de medios de transporte, la gestión y regulación de medios de transporte público privado, el uso de bicicletas para desplazamientos de trayectos medianos e integración de soluciones como el “último kilómetro”, el cual permite a ciudadanos tomar en préstamo bicicletas para desplazarse en trayectos cortos y medios, que puedan adaptarse, cambiar y reducir en embotellamiento del tráfico en algunos sectores críticos en la ciudad.

En la Figura 1 se presentan algunos tipos de movilidad como la bicicleta, el transporte masivo llamado Transmilenio; el porcentaje de viajes realizados por estrato, el tipo de género o sexo y el motivo del viaje, como compras, estudio y regreso a casa, indicadores relevantes en la ciudad de Bogotá.

Figura 1. Algunos indicadores de movilidad en Bogotá



Fuente: Secretaría Distrital de Movilidad (2019).

### Elementos para iniciar una transformación digital

Según Chowdhury *et al.* (2017), se pueden considerar como elementos relevantes en la iniciativa

de la transformación digital los que se visualizan en la Tabla 1. Entre ellos están la concientización, los consumidores exigentes, el modelo de negocio, la plataforma digital, el internet de las cosas, el Big Data y la tecnología digital.

Tabla 1. Elementos para iniciar una transformación digital

Elementos	Proceder
Concientización	Sensibiliza a empresarios, consumidores y Gobierno.
Consumidores exigentes	Empiezan a vivir en un entorno tecnológico, bajo una transformación digital
Modelo de negocio	Migra sus procesos de negocio centralizados a un modelo global y flexible
Plataforma digital	La nueva dinámica social y económica debe ser atendida por las instituciones, requiere de una transformación que entienda el nuevo lenguaje de los consumidores
Internet de las cosas	El intercambio de la información entre dispositivos es una de las principales áreas que hay que trabajar para mantener una comunicación constante y en tiempo real. Esto permitirá satisfacer las nuevas exigencias de los consumidores actuales
Big Data	El volumen "masivo" de los datos debe ser explotado al máximo para identificar las situaciones actuales reales
Legalización de la tecnología abierta y divisas digitales	Aplicación de sistemas y políticas de seguridad que salvaguarden los intereses de los empresarios y consumidores

Fuente: Chowdhury *et al.*, (2017).

De acuerdo con el contexto anterior, es relevante identificar la relación que existe entre la generación de las últimas tecnologías como es el Big Data, con el desarrollo urbano sostenible en la ciudad de Bogotá. Esto, a partir de la interacción de la mayoría de los dispositivos y la comunicación con infraestructura, procesamiento de grandes cantidades de datos, lo que permite un encaminamiento directo a un ITS en esta ciudad (Aristizábal, 2016).

### El Big Data desde los ITS

Siguiendo la recomendación de Rocha y Villareal (2018), se aplica la técnica del Big Data enfocada en tres principios: 1) Análisis de los datos. Tomando aquellos datos relevantes y dejando en una bitácora los que por el momento no aportan para el tema específico;

2) generar valor agregado en los datos seleccionados no redundantes en el principio anterior. Este valor agregado se estipula con una serie de preguntas que, al procesar los datos, se contestan y aportan significativamente al tema de estudio; y por último, 3) el diseño de una prueba piloto para visualizar los resultados obtenidos, logrando su análisis y evaluación y concluir al respecto.

### Beneficios del Big Data

En la Tabla 2 se acopian algunas definiciones importantes como apoyo a esta herramienta BD de última generación bajo la industria de 4.0, el cual se caracteriza por su velocidad, variedad y veracidad en el uso y manejo de los datos estructurados y no estructurado al momento de resolver la problemática requerida.

Tabla 2. Algunas definiciones relevantes del Big Data

Autores	Proceder
Aportes	Sensibiliza a empresarios, consumidores y Gobierno.
Serrano Cobos (2013)	Big Data es el término que se usa cada vez más para describir la técnica para recopilar y analizar datos "masivos" estructurados y no estructurados para dar soluciones complejas.
Vásquez García (2015)	Big Data es un conjunto de técnicas informáticas que sirven para almacenar, procesar y gestionar volúmenes de información. Un sistema de Big Data debe ser veloz, tener variedad en los datos que almacena, es decir, ser capaz de guardar tipos de datos de diferentes estructuradas.
Rogelez <i>et al.</i> (2013)	Big Data es un término que se aplica a un conjunto de datos cuyo tamaño o tipo está más allá de la capacidad de las bases de datos tradicionales para capturar, administrar y procesar datos con baja latencia. De igual forma, presenta algunas características como alto volumen, alta velocidad, alta variabilidad y alta veracidad.
Tabares y Hernández (2014)	Big Data permite la disponibilidad de grandes cantidades de información estructuras y desestructurada en tiempo real, por lo tanto, la dimensión clave de dicha terminología se concentra en el volumen, variedad y velocidad en sus datos.

En la Figura 2 se visualiza la caracterización del BD desde el modelo de 4V, (volumen, velocidad, valor y veracidad).

**Figura 2. Caracterización del Big Data desde el modelo de 4Vs**

#### **VOLUMEN**

Es la característica más asociada al Big Data, pues su propia denominación lo indica. Es el volumen de datos que se extra, almacena y procesa.

#### **VELOCIDAD**

La palabra más indicada en la actualidad para esta característica es el denominado real time (tiempo real), ya que la información fluye a gran velocidad y debe procesarse de manera rápida y oportuna.

#### **VARIEDAD**

Los datos pueden ser estructurados o no estructurados, además de provenir de distintas fuentes y formatos.

#### **VERACIDAD**

Es la característica que representa la calidad de los datos. Se busca evitar al máximo la contaminación de los mismos por causa del ruido.

Fuente: Chaga y Ferraz (2017).

## **Metodología**

Esta investigación tiene un enfoque de perspectiva cualitativa de tipo descriptivo (Picón, 2019), el cual facilita la comprensión e interpretación de los significados, las concepciones, experiencias y prácticas alrededor de un fenómeno identificado, por medio de una inmersión en el entorno real. Esta inmersión se realiza en dos partes: primero, desde los autores con estudios y proyectos del Big Data en los ITS, consultas en bases de datos seguras y, segundo, por medio del análisis de la estructura de datos recogidos, que permiten identificar el estado de los datos previamente agrupados de acuerdo con una serie de indicadores relacionados con el tema de estudio. Estos elementos permitirán un primer acercamiento para conocer la magnitud, importancia y veracidad que podría ofrecer un Big Data como componente fundamental en el seguimiento

y prospectiva de un ITS para Bogotá, el cual vienen desarrollando desde el año 2015 entes gubernamentales como la Secretaría Distrital de Movilidad [SDM] y la Empresa de Telecomunicaciones de Bogotá [ETB], entre otros.

### **Técnicas de recolección y análisis de datos**

Se trabaja con un conjunto de datos, las cuales se organizan en línea con las necesidades del estudio, donde reflejan elementos primordiales del transporte público y privado, al igual que el sentir de los usuarios del transporte en Bogotá. La fuente principal de estos datos corresponde al SDM en Bogotá durante el periodo 2016-2019. Es importante resaltar que en esta etapa I de la investigación no se analiza el contenido de los datos y su accionar, sino su estructura, integridad y redundancia de estos, contruidos por el SDM.

## Procedimiento

Inicialmente, se da una fase exploratoria de búsqueda de información con autores que aportan en la proyección al ITS en bases de datos seguras y luego se hace una búsqueda de datos libres desde la SDM con el fin de analizar y evaluar su estructura, integridad y redundancia en su contenido.

En segunda instancia, se lleva a cabo la fase de indagación contactando a los coordinadores de las entidades SDM y ETB correspondiente al proyecto ITS, por medio de correos institucionales, con miras a la socialización de la propuesta y generar el interés, que se ve reflejado en la aceptación del consentimiento informado. De igual forma se exponen elementos y componentes fundamentales que aportan a la propuesta del Big Data para el seguimiento en la proyección del ITS en Bogotá.

En tercera instancia se hace un diagnóstico desde la zona de estudio para establecer puntos fuertes y débiles de los ITS. Por último, como cuarta instancia, se visualiza el panorama del contexto con las TIC y la aplicación de los principios del Big Data hacia las tendencias de la situación.

## Resultados y discusión

### Algunos referentes desde el Big Data

Se aporta desde la tecnología, innovación, arquitecturas y algoritmos, entre otros, citando fuentes nacionales e internacionales publicadas durante el periodo 2013 al 2019, como se observa en la Tabla 3. Se resaltan estudios y proyectos realizados con el aporte de las técnicas del Big Data.

**Tabla 3. Trazabilidad del uso del Big Data**

Línea de tiempo	Aportes
2013-2015	Metodologías V2V de vehículo a vehículo (Hafner <i>et al.</i> , 2013). D2D de dispositivo a dispositivo o de celular a celular (Chen <i>et al.</i> , 2015). VANET para agrupación vehicular (Lv <i>et al.</i> , 2015). BD desde la nube. De igual forma, Djahel <i>et al.</i> (2016) con el proyecto MAS basado en una guía para identificar los cambios en el tiempo de movilización en el transporte de los vehículos a distancias cortas y largas.
2016	RNA's (Wang <i>et al.</i> , 2016); (Cao <i>et al.</i> , 2020) y Heurístico (Wu & Tan, 2016); (Gruntzki & Bazza, 2106).
2017	Deep Learning para el aprendizaje automático de patrones de tráfico (Zanoon <i>et al.</i> , 2017). Uso de dispositivos inteligentes, por ejemplo: tarjetas de pasajeros (Xiao <i>et al.</i> , 2017). Empleo de nuevos algoritmos de inteligencia artificial avanzados, como el de identificación de densidad de Kernel (Jin, 2017).
2018	Tendencia: usar como principal fuente de formación las redes sociales (Chen <i>et al.</i> , 2018). Uso de la nube y la clusterización con Hadoop (Massobrio <i>et al.</i> , 2018).
2019	Zhu <i>et al.</i> (2019) resaltan el uso de las metodologías adaptativas (se recalibra y ajusta de manera automática) y temporales para la predicción de flujos y eventos vehiculares (Asadi <i>et al.</i> , 2019). De igual forma, algoritmos de inteligencia artificial, plataformas de aprendizaje y genéticos combinados con simulación (Tian <i>et al.</i> , 2019)

Hafner *et al.* (2013) y Lv *et al.* (2015) parten de un estudio con un enfoque del flujo de tráfico *versus* Big Data y utilizan en común métodos de predicción de flujo de tráfico basados en aprendizaje profundo, donde se consideran correlaciones espaciales y temporales de manera inherente. Mientras que Djahel *et al.* (2015) y Wang *et al.* (2016) se basan en un método arquitectónico de control robótico para el seguimiento continuo a los cruceros en tiempo real de tráfico, desde la predicción en el control y valor de su velocidad, aceleración, frenos y otros.

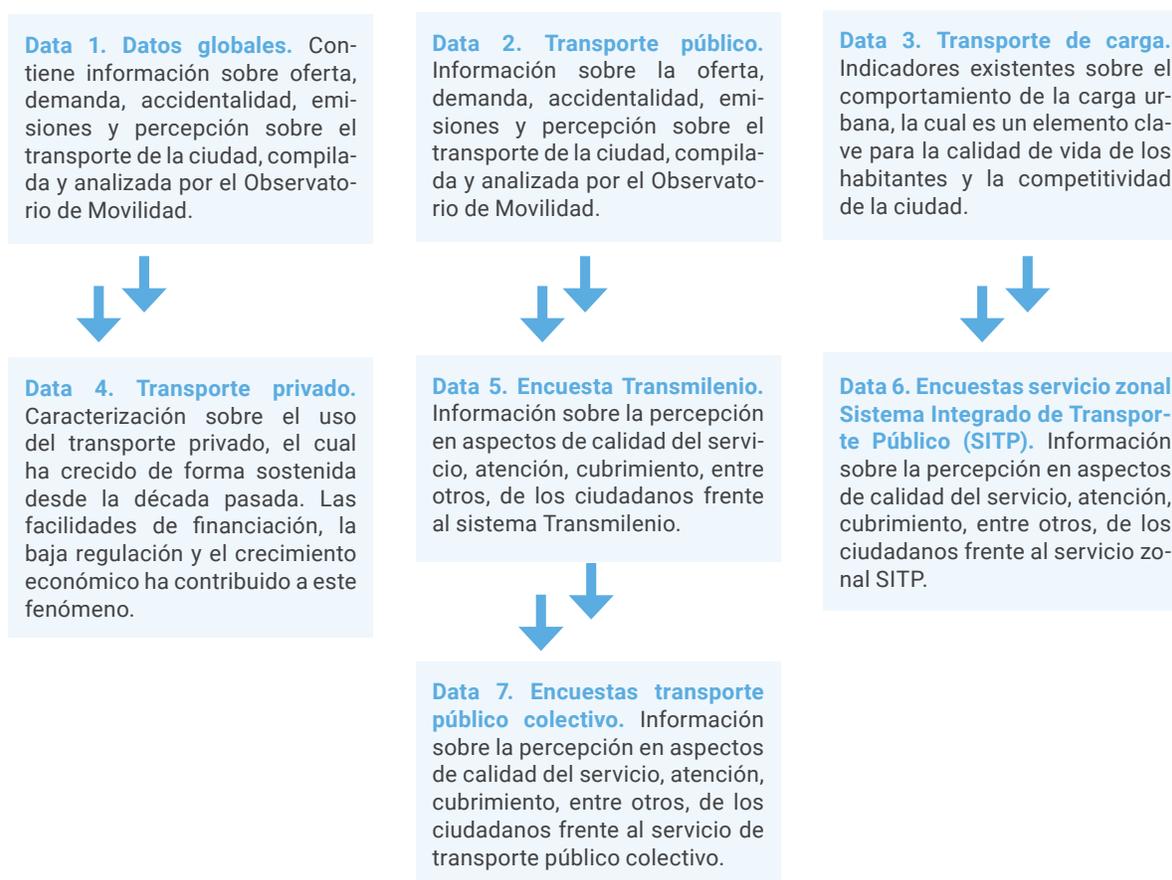
De igual forma, en un estudio de las ciudades inteligentes Wu y Tan (2016), Xiao *et al.* (2017) y Jin (2017) promocionan una urbanización sostenible, proponen un marco jerárquico de control y gestión para fusión de tecnología centralizada *top-Down* y distribuida, basada en humanos *bottom-up*, para hacer que las infraestructuras urbanas sean inteligentes y seguras. Reconoce el origen de los datos, prioridad, vivacidad y calidad y pone a disposición para análisis posterior, es posible implementar computación en la nube para facilitar minería de datos y análisis predictivo.

Por último, Chen *et al.* (2018), Tian *et al.* (2019) y Zhu *et al.* (2019) parten del uso de las técnicas de detección de vehículos para sistemas anticolidión. Estos autores plantean un diseño de sistemas inteligentes, con una cuidadosa selección de sensores y detección para lo cual se requieren algoritmos. Aquí, proporcionamos discusión, críticas y perspectiva sobre sensores, detección de vehículos, seguimiento, sistemas de detección de motocicletas y detección nocturna de vehículos.

### Estado de las estructuras de DATAS

En la Figura 3 se inicia con siete grupos de datos analizados desde su contenido en estructura, integridad y redundancia como posibles componentes de aporte a la proyección del ITS para Bogotá. Se parte de datos globales, indicadores de demanda accidental, comportamiento de la carga urbana, caracterización sobre el uso del transporte privado, datos de sentimientos como la percepción de calidad de los servicios, cubrimientos y aspectos del transporte público colectivo.

**Figura 3. Datos estructurados como aporte al Big Data**



Fuente: Autores.

En la Tabla 4 se aprecia el análisis realizado del contenido de los datos descritos en la tabla anterior.

**Tabla 4. Estado del contenido de las siete DATAS**

DATAS	Total Registro	Total variables	Estado del contenido		
			Redundancia	Integridad	Vacío (null)
1	143 122	33	45 %	3 %	84 %
2	24 831	70	28 %	5 %	71 %
3	185 125	27	25 %	2 %	89 %
4	25 123	44	3 %	4 %	56 %
5	24 123	20	18 %	15 %	34 %
6	12 456	45	21 %	12 %	43 %
7	14 672	84	15 %	16 %	52 %

Fuente: Open Data – SDM (2019).

Llama la atención en este grupo de datos tomados al azar, el estado de su contenido desde la redundancia, integridad y nulidad en los datos. De ahí la importancia de validar de manera significativa los datos libres que se encuentren en internet, desde su confiabilidad durante el proceso de predicción para la toma de decisiones. Examinando los datos recabados a lo largo de la investigación, sale a la luz las innumerables oportunidades y desafíos para la propuesta de un Big Data que apoye la proyección del ITS en Bogotá. Además, se identifican debilidades del sistema al mejorar la gobernabilidad en el procesamiento efectivo de datos, aplicación de la inteligencia artificial, la analítica de datos y otras áreas del conocimiento.

### **Se debe mejorar solo lo que se puede medir**

Se parte desde el principio del valor agregado medible al lograr tres grandes retos para impactar al sector público con las técnicas del Big Data: 1) políticas desde la ética; 2) gobernanza en el uso y manejo de los datos y 3) capacitar a los directivos, colaboradores y todos los actores del contexto.

### **Uso y protección de datos personales**

Al analizar el conjunto de datos de la SDM, se visualiza un riesgo primordial como es la privacidad de los datos. Como afirma Villalobos (2020), es un derecho de los sujetos en la

sociedad el de no ser observado y en especial, respetar su privacidad. Partiendo de la gran cantidad de dispositivos digitales que se encuentran en el entorno y segundo a segundo producen millones de datos no autorizados para ser divulgados, datos que son esenciales en la técnica del Big Data y la IA. Por tanto, es fundamental utilizar mecanismos donde las personas no se sientan observadas y se cumpla la normativa y regulación en la protección de datos en cada país.

### **Plan de acción desde el aporte del Big Data**

Se recomienda partir de un centro especial de alto volumen de datos, con tecnología de software y hardware adecuada, segura y protegida contra todo tipo de riesgo, en especial del ataque de los hackers. En la Tabla 5 se muestran algunos requisitos desde su arquitectura, operatividad, fiabilidad e inversión, que garanticen en gran parte la versatilidad y flexibilidad del Big Data propuesto. Estos elementos se conocen como clasificación TIER llamada capa o nivel, siendo un tipo de certificación que pretende avalar el desempeño y la confiabilidad de las infraestructuras de los Data Centers. Está dividida en cuatro categorías: TIER I, TIER II, TIER III y TIER IV. En cada nivel va midiendo la usabilidad y accesibilidad, al igual que la operatividad, versatilidad, tolerancia a falla, y algo fundamental: la facilidad para la recuperación de desastres naturales (Piguave, 2022).

**Tabla 5. Apropiación de las categorías tier para el centro de Big Data propuesto**

Línea de tiempo	Aportes
TER I	El nivel inicial posee componentes no redundantes y una única vía, también no redundante para la distribución de energía. De esta forma, las fallas en los componentes como servidores, dispositivos de telecomunicaciones, sistemas de almacenamientos o equipos de enfriamiento pueden impactar de forma negativa en los sistemas.
TER II	En esta categoría, ciertos elementos que componen el sistema son redundantes. Los servidores pueden encontrarse conectados a un sistema de alimentación interrumpida (UPS), donde existe la posibilidad que la infraestructura sea paralizada por determinados eventos, tanto planificados como aquellos que ocurren por fuerza mayor.
TER III	Puede considerarse un servicio de almacenamiento interrumpido. Cuenta con componentes redundantes y dos fuentes de energía, una activa y otra alterna. A pesar de estas cualidades, los periodos de mantenimiento pueden generar algunas breves interrupciones, que serán atendidas gracias a la línea de energía secundaria mitigando el riesgo.
TER IV	Es la categoría más elevada en cuanto a la confiabilidad. Incorpora múltiples fuentes de energías redundantes y un sistema de refrigeración sofisticados. A su vez, debido a sus diversas fuentes de energía, es posible realizar labores de mantenimiento sin presentar interrupciones y está preparado para siniestros como: terremotos, incendios y fallos eléctricos.

Por tanto, la predicción de estos datos tiende a ser proporcional más el doble y es allí cuando es necesario reevaluar periódicamente la adquisición periódica de la actualización de los equipos hasta lograr la categoría aceptable según la distribución propuesta en la Tabla 5 (Massobrio *et al.*, 2018).

### **Carga de datos**

Este proceso se realiza en un “bajo tiempo” debido al requerimiento de un “mejor tiempo” de respuesta a la operación solicitada. Es aquí cuando es primordial pensar que al “expandirse el sistema” en su proyección, no requiere actualizarse por completo debido a la eficiencia en tiempo y costos, sino actualizar los componentes necesarios para irse acoplado a esta dinámica futurista.

### **Ubicación del centro de datos - Big Data**

El sitio seleccionado debe cumplir con los requisitos ambientales pertinentes, bloqueo a intrusos o desconocidos en el sistema, creación de un sistema de credenciales protegido, contar con personal experto y autorizado en el sistema, en especial de ciberseguridad, redes y conectividad.

### **Técnicas desde el Big Data propuesto**

La aplicación de estas técnicas requiere de personal altamente calificado en procesos de analítica de datos, ciencia de datos, machine learning e IA (Aristizábal, 2016), debido a que los requerimientos obligan, en la mayoría de los casos, a la interacción con algoritmos

especializados y herramientas para obtener y transmitir los resultados esperados, no sin antes revisar y analizar cuidadosamente los “complementarios”; es decir, evaluar que la relación de valores de variables, prioridades y categorías sean las adecuadas para una predicción cercana a la realidad. Es por ello que debe existir una comunicación permanente y continua entre todos los integrantes del equipo de trabajo de acuerdo con sus funciones y responsabilidades.

### Aportes en áreas de oportunidad

Los aportes obtenidos en este estudio parten de un área de oportunidad desde la instrumentación de lugares o sitios de embotellamiento de tráfico, por medio de métodos y acciones al analizar datos “masivos” desde la SDM en intervalos de años con datos estructurados tomados con los mismos transportistas, lo cual ofrece la facilidad para predecir bajo la técnica del Big Data sobre el seguimiento en el tráfico (Asadi y Regan, 2019). Al lograr establecer algoritmos de aprendizaje en un modelo específico como apoyo a estos mismos para la reprogramación en tiempo real de su partida y llegada para evitar algún riesgo en el camino (Ríos, 2020).

No cabe duda de que el Big Data, en palabras de Barba (2018), ha venido a revolucionar la manera de vivir de las personas, aprovechando al máximo otras herramientas de última generación como es el IoT, así como el uso y manejo de celulares inteligentes, como otros elementos fundamentales como Arduino y Raspberry (Djahel *et al.*, 2015). Son dispositivos fáciles de instalar, actualizar y programar con algoritmos de aprendizaje específicos en tiempos muy cortos, produciendo datos que

continuamente alimentarán las predicciones del Big Data propuesto.

En la etapa II de esta investigación se espera presentar el grupo de datos desde los principios del Big Data, y generar predicciones que aporten a la proyección del ITS con datos estructurados y no estructurados desde la inteligencia artificial, con el uso de herramientas como Python, R, Power Bi y otras.

### Conclusiones

El desafío de un Big Data en Bogotá, bajo un centro de acopio, es un proceso bastante complejo. Es responsabilidad de los entes gubernamentales SDM, ETB, la academia y otros, contribuir de forma positiva y proactiva a la propuesta de un ITS que controle el transporte público y privado en una ciudad que día tras día crece, no solo en su población sino en la necesidad de contar con servicios oportunos, seguros y confiables.

De igual forma, el procesamiento electrónico de los datos en tiempo real en el transporte público y privado por medio de tecnologías de última generación como el Big Data, *Machine Learning* y la IA ha potenciado las aplicaciones y recursos de los ITS. No puede dejarse de lado la importancia de la seguridad y protección de este ITS, no solo en cuanto al resguardo de su integridad sino su operatividad, versatilidad y flexibilidad en las etapas de recuperación, reiniciación y puesta en marcha, para evitar que ocasionen vulnerabilidades en el sistema.

Los aportes obtenidos en esta etapa I del estudio, no solo parten desde la literatura con proyectos de inicio, avance y terminados en

Latinoamérica, Norteamérica, Europa y Asia, ilustrando tecnologías, modelos, métodos, técnicas y procedimientos que fortalecen los a ITS en las ciudades grandes, sino del análisis de la estructura y esquema que presenta el contenido del grupo de datos recolectados en la ciudad de Bogotá.

Este contenido señala indicadores como: datos globales, demanda accidental, comportamiento de la carga urbana, caracterización sobre el uso del transporte privado, la percepción de calidad del servicios, cubrimientos y aspectos del transporte público colectivo entre otros, generados por los mismos transportistas por medio de sus dispositivos y periféricos electrónicos, visualizando con gran preocupación que todas presentan un alto porcentaje de errores en su contenido revisando desde la redundancia, integridad y nulidad.

## Recomendaciones

Es importante revisar el sitio web o la ruta que se selecciona para descargar “datos libres”, la cual debe ser confiable y segura, de lo contrario no se garantiza el uso y manejo de estos datos.

Evitar “datos libres” que se encuentran en repositorios desconocidos o llegan en la nube. El desconocer su procedencia, afectaría la predicción y tendencias en una situación real.

Los datos libres de la ciudad de Bogotá son limitados y lo más preocupante es que la gran mayoría de estos datos no reflejan la situación actual que se tiene en la ciudad, es decir, no hay dinámicos en estos. De ahí, la importancia de evaluar primero la calidad de los datos que

se recogen para lograr determinar que tan relevante son para dar veracidad al inicio de una investigación.

## Referencias

- **Álvarez, A., & Patiño, A.** (2015). Arquitectura empresarial y metodologías ágiles-Una combinación efectiva para hacer frente a los frecuentes cambios en el negocio. *Actas de Ingeniería, 1*, 145-152. <https://es.scribd.com/document/410099179/Articulo-3-Arquitectura-empresarial-y-metodologias-agiles-Una-combinacion-efectiva-para-hacer-frente-a-los-frecuentes-cambios-en-el-negocio-pdf#>
- **Arango-Serna, M. D., Londoño-Salazar, J. E., & Branch-Bedoya, J. W.** (2015). Enfoque de arquitectura de solución, mecanismo para reducir la brecha entre la arquitectura empresarial y la implementación de soluciones tecnológicas. *Dyna, 82*(193), 117-126. <http://dx.doi.org/10.15446/dyna.v82n193.46752>
- **Aristizábal, J. A.** (2016). Analítica de datos de aprendizaje (ADA) y gestión educativa. *Gestión de la educación, 6*(2), 149-168. <https://doi.org/10.15517/rge.v1i2.25499>
- **Asadi, R., & Regan, A.** (2019). A Spatial-Temporal Decomposition Based Deep Neural Network for Time Series Forecasting. *ArXiv, 1*(1), 1-17. <http://arxiv.org/abs/1902.00636>
- **Barba González, C.** (2018). *Big Data Optimization: Algorithmic Framework for Data Analysis Guided by Semantics*. [Tesis doctoral Universidad de Málaga]. <https://riuma.uma.es/xmlui/handle/10630/17264?show=full>
- **Bbvaapimarket.com.** (2020). *Reconocer los diferentes tipos de datos, indispensable la era del Big Data*. <https://www.bbvaapimarket.com/es/mundo-api/reconocer-los-diferentes-tipos-de-datos-indispensable-en-la-era-del-big-data/>

- **Cao, T., Matyas, J. J., Renn, C. L., Faden, A. I., Dorsey, S. G., & Wu, J.** (2020). Function and mechanisms of truncated BDNF receptor TrkB. T1 in neuropathic pain. *Cells*, 9(5), 1194. <https://doi.org/10.3390/cells9051194>
- **Cacho-Elizondo, S., & Lázaro Álvarez, J. D.** (2018). Transformación digital de los negocios de network marketing, multinivel y ventas directas. El impacto de nuevas herramientas y plataformas digitales. *The Anáhuac Journal*, 18(2), 13-41.
- **Chagas, A., & Ferraz, C.** (2017). *Smart Vehicles for Smarter Cities: ContextAware V2X to Improve Qol*. Anais Do XXIII Simpósio Brasileiro de Sistemas. <https://www.inf.ufgrs.br/webmedia2017/wp-content/anaiswebmedia/files/wtd/wtd11.pdf>
- **Chen, Y., Lv, Y., Wang, X., Li, L., & Wang, F. Y.** (2018). Detecting Traffic Information from Social Media Texts With Deep Learning Approaches. *IEEE Transactions on Intelligent Transportation Systems*, 1(1), 1–1. <https://doi.org/10.1109/TITS.2018.2871269>
- **Cheng, J., Cheng, J., Zhou, M., Liu, F., Gao, S., & Liu, C.** (2015). Routing in internet of vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 16(5), 2339–2352. <https://doi.org/10.1109/TITS.2015.2423667>
- **Chowdhury, M., Apon, A., & Dey, K.** (2017). *Data Analytics for Intelligent Transportation Systems*. Elsevier. <https://www.sciencedirect.com/book/9780128097151/data-analytics-for-intelligent-transportation-systems>
- **International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, China, 2017**, pp. 315-320, <https://doi.org/10.1109/CSE-EUC.2017.63>
- **Djahel, S., Jabeur, N., Barrett, R., & Murphy, J.** (2015). *Toward V2I communication technology-based solution for reducing road traffic congestion in smart cities*. International Symposium on Networks, Computers and Communications, ISNCC 2015, 1(1), 6. <https://doi.org/10.1109/ISNCC.2015.7238584>
- **Hafner, M. R., Cunningham, D., Caminiti, L., & Del Vecchio, D.** (2013). Cooperative collision avoidance at intersections: Algorithms and experiments. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1162–1175. <https://doi.org/10.1109/TITS.2013.2252901>
- **Jin, Y., Li, J., Ma, D., Guo, X., & Yu, H.** (2017). *A Semi-Automatic Annotation Technology for Traffic Scene Image Labeling Based on Deep Learning Preprocessing*. IEEE International Conference on Computational Science and Engineering (CSE) and IEEE
- **Kitzmann, N., Romanczuk, P., Wunderling, N., & Donges, J. F.** (2022). Detecting contagious spreading of urban innovations on the global city network. *European Physical Journal - Special Topics*, 231(9), 1609-1624. <https://doi.org/10.1140/epjs/s11734-022-00470-4>
- **Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y.** (2015). Traffic Flow Prediction with Big Data: A Deep Learning Approach. *IEEE Transactions on Intelligent Transportation*. 16(2), 885-873. <https://doi.org/10.1109/TITS.2014.234566>
- **Molina Araque, F. A., & Espitia Pineda, A. A.** (2016). *Propuesta de diseño e implementación de una aplicación móvil (App) como plataforma para compartir el uso del carro particular (carpooling) entre empleados de las sucursales en Bogotá de la empresa Claro Colombia SA., para teléfonos inteligentes cuyo sistema operativo sea Android 5.1* [tesis de especialización. Universidad Distrital Francisco José de Caldas]. <https://repository.udistrital.edu.co/handle/11349/4536>
- **Massobrio, R., Nesmachnow, S., Tchernykh, A., Avetisyan, A., & Radchenko, G.** (2018). Towards a Cloud Computing Paradigm for Big Data Analysis in Smart Cities. *Programming and Computer Software*, 44(3), 181–189. <https://doi.org/10.1134/S0361768818030052>

- **Moovit.** (2022). *Estadísticas de transporte público en Bogotá*. [https://moovitapp.com/insights/es/Moovit\\_Insights\\_%C3%8Dndice\\_de\\_Transporte\\_P%C3%BAblico\\_Colombia\\_Bogota-762](https://moovitapp.com/insights/es/Moovit_Insights_%C3%8Dndice_de_Transporte_P%C3%BAblico_Colombia_Bogota-762)
- **Piguave** Peralta, D. (2022). Estudio de factibilidad de una Data Center para el procesamiento de la información en la carrera de Tecnologías de la Información [tesis grado, Universidad Estatal del Sur de Manabí] [https://publications.pik-potsdam.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item\\_26817](https://publications.pik-potsdam.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_26817). <http://repositorio.unesum.edu.ec/handle/53000/4321>
- **Rincón-García**, N., Bamford, D., Mayers, M., Carreño Vargas, J. E., & Roldan Nariño, R. F. (2020). Herramientas de transformación digital para mejorar la planificación urbana mediante el uso de la metodología de caso de negocio de transporte. *Revista Espacios*, 41(31), 296-304. <https://www.revistaespacios.com/a20v41n31/a20v41n31p24.pdf>
- **Ríos** Rodríguez, L. E. (2020). *Análisis de datos de sentimientos enfocados al servicio de transporte masivo Transmilenio SA aplicando tecnologías Big Data*. <http://hdl.handle.net/11371/2907>
- **Rocha** Tamayo, E., & Villarreal Vásquez, C. (2018). *Impacto del uso de Big Data en la gestión de movilidad vehicular en el cantón de la libertad* [Tesis de grado, Universidad de Guayaquil]. <http://repositorio.ug.edu.ec/handle/redug/27297>
- **Rogelez**, F. J., Molina Hernández, I., & Marín, A. M. (2013). Sistema de semaforización inteligente en la ciudad de Bogotá para mejorar los tiempos de recorrido del sistema Transmilenio. *Traza*, 1(7), 10-29. <https://ciencia.lasalle.edu.co/tr/vol1/iss7/2/>
- **Secretaría** Distrital de Movilidad – SDM. (2019). *Encuesta de movilidad en Bogotá*. [https://observatorio.movilidadbogota.gov.co/encuestas\\_de\\_movilidad](https://observatorio.movilidadbogota.gov.co/encuestas_de_movilidad)
- **Serrano-Cobos**, J. (2013). Microsoft Big Data y not so Big Data. *Anuario ThinkEPI*, 7, 161-163.
- **Suat-Rojas** N. (2021). *Extracción y análisis de información de accidentes de tránsito desde redes sociales*. <http://repositorio.unal.edu.co/handle/unal/80927>
- **Tabares**, L. F., & Hernández, J. F. (2014). *ISAM. Big Data Analytics: Oportunidades, Retos y Tendencias*. Universidad de San Buenaventura. <https://gc.scalahed.com/recursos/files/r161r/w25569w/Big%20Data%20Analytics.pdf>
- **Tian**, Y., Hu, W., Du, B., Hu, S., Nie, C., & Zhang, C. (2019). IQGA: A Route Selection Method Based on Quantum Genetic Algorithm - Toward Urban Traffic Management Under Big Data Environment. *World Wide Web*, 22(5), 2129–2151. <https://doi.org/10.1007/s11280-018-0594-x>
- **Vásquez** Ardila J., M. (20 de julio del 2021.) *National Transportation Systems Center. Tránsito de Bogotá tomará decisiones a partir de datos, tecnología e innovación*. <https://bogota.gov.co/mi-ciudad/movilidad/tecnologia-e-innovacion-nuevo-aliado-del-transito-de-bogota>
- **Vásquez** García, D. M. (2015). *Diseño de la Arquitectura Empresarial para el Sistema Inteligente de Transporte de Bogotá*. Oracle. <https://repository.udistrital.edu.co/handle/11349/2378>
- **Villalobos** Fonseca, H. (2020). O desenvolvimento tecnológico em matéria policial: uma receita de sucesso para prevenir o delito. *Revista de Relaciones Internacionales, Estrategia y Seguridad*, 15(1), 79-97. <https://doi.org/10.18359/ries.4243>
- **Wang**, S., Djahel, S., Zhang, Z., & McManis, J. (2016). Next Road Rerouting: A Multiagent System for Mitigating Unexpected Urban Traffic Congestion. *IEEE Transactions on Intelligent Transportation Systems*, 17(10), 2888–2899. <https://doi.org/10.1109/TITS.2016.2531425>

- **Wu, Y., & Tan, H.** (2016). Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework. *School of Mechanical Engineering, Beijing Institute of Technology, 1(1)*, 1–14.: <https://doi.org/10.48550/arXiv.1612.01022>
- **Xiao, Z., Ponnambalam, L., Fu, X., & Zhang, W.** (2017). Maritime Traffic Probabilistic Forecasting Based on Vessels' Waterway Patterns and Motion Behaviors. *IEEE Transactions on Intelligent Transportation Systems, 18(11)*, 3122–3134. <https://doi.org/10.1109/TITS.2017.2681810>
- **Zanoon, N., Al-Haj, A., & Khwaldeh, S. M.** (2017). Cloud Computing and Big Data is there a Relation between the Two: A Study. *International Journal of Applied Engineering Research, 12(17)*, 973–4562. <http://www.ripublication.com>
- **Zhu, M., Chen, W., Xia, J., Ma, Y., Zhang, Y., Luo, Y., Zhaosong, H., & Liu, L.** (2019). Location2vec: a situation-aware representation for visual exploration of urban locations. *IEEE Transactions On Intelligent Transportation Systems, 20(10)*, 3981-3990. <https://doi.org/10.1109/TITS.2019.2901117>