



Fundación Universitaria
SAN MATEO

TECNOLOGÍA EN DESARROLLO DE SOFTWARE



Fundación Universitaria
SAN MATEO

FACULTAD DE INGENIERÍA Y AFINES
PROGRAMA DE INGENIERÍA DE SISTEMAS

MANEJO DE BIG DATA EN LAS REDES SOCIALES
TRABAJO DE GRADO MODALIDAD DE OPCIÓN DE GRADO

JORDAN STEVEN BENAVIDES ACOSTA
EDWARD YEZID RUIZ PUENTES

DIRECTOR(A)
JOSE FERNANDO LÓPEZ QUINTERO

BOGOTÁ
2022

NOTA DE SALVEDAD DE RESPONSABILIDAD INSTITUCIONAL

“La Fundación Universitaria San Mateo NO se hace responsable de los conceptos emitidos en el presente documento, el departamento de investigaciones velará por el rigor metodológico de la investigación”.

CONTENIDO

INTRODUCCIÓN	7
CAPÍTULO I: DESCRIPCIÓN DEL PROYECTO	1
Presentación del problema de investigación	1
Justificación	3
<i>Objetivo general</i>	5
<i>Objetivos específicos</i>	5
CAPÍTULO II: MARCO TEÓRICO	6
Antecedentes de la investigación	6
Bases teóricas o fundamentos conceptuales.....	13
Bases legales de la investigación	15
CAPÍTULO III: DISEÑO METODOLÓGICO	20
Tipo de investigación.....	20
Población	21
Técnicas e instrumentos de recolección de datos	21
CAPÍTULO IV: RESULTADOS DE LA INVESTIGACIÓN	22
Resultados del objetivo específico no. 1	22
Resultados del objetivo específico no. 2	22
Resultados del objetivo específico no. 3	22
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES	23
BIBLIOGRAFÍA	23

RESUMEN

En los tiempos de hoy día el mundo está bajo un sistema masificado y globalizado por el internet, y teniendo en cuenta lo anterior no se tardó mucho en desarrollar las redes sociales como método de comunicación, ya sea de manera escrita y/o audiovisual. Donde, desde el principio se ha podido percatar la cantidad de usuarios que buscan hacer uso de estas herramientas de comunicación generando una gran demanda de recursos a las empresas propietarias de dichas redes sociales. Ya teniendo este contexto en mente gran parte de las compañías ha empezado hacer uso del Big Data para cualquier contexto, ya sea con fines investigativos, en el caso de muchos de los avances en temas de enfermedades y/o problemas ambientales, o también para mercadotecnia como sería el caso de grandes empresas como Amazon, eBay entre otras, y con esta investigación se busca saber cómo las redes sociales hacen el manejo del Big Data.

PALABRAS CLAVE:

Big Data; Redes sociales; Manejo de datos; Gestión de archivos; Digitalización.

ABSTRACT

In today's world the world is under a massive and globalized system by the internet, and taking into account the above it did not take long to develop social networks as a method of communication, either written and / or audiovisual. Where, from the beginning it has been possible to notice the number of users seeking to make use of these communication

tools generating a great demand for resources to the companies that own these social networks. With this context in mind, most companies have started to make use of Big Data

KEYWORDS:

Big Data; Social media; Data handling; Management archives; Digitization

INTRODUCCIÓN

La cantidad de datos al inicio de las redes sociales era muy pequeña ya que no todos tenían alcance a estas, pero a medida del paso del tiempo más, y más personas tiene alcance a las redes sociales, con lo cual de manera proporcional la cantidad de datos a incrementado.

El Big Data puede ser de gran ayuda, siempre y cuando la información y/o datos sean usados de manera correcta. Con base a lo anterior “nace” esta investigación ya que se busca saber cómo es el manejo de una gran cantidad de datos (Big Data) que realizan las redes sociales.

Lo que se busca plantear con esta investigación es definir el proceso para el estudio del Big Data en las redes sociales, para después empezar a profundizar los métodos que se usan para su aplicación.

Describiendo su contenido en capítulos de la siguiente forma:

- Capítulo 1
 - Contexto de la situación
 - ¿Qué es Big data?
 - ¿Qué es Hadoop?
 - Uso de apache hadoop
 - Cómo se aplica en el marketing
- Capítulo 2
 - Antecedentes
 - Bases teóricas
 - Bases legales
- Capítulo 3
 - Tipo de investigación
 - Población
 - Técnicas e instrumentos de investigación
- Capítulo 4

- Resultados de la investigación
- Capítulo 5
 - Conclusiones

Junto a lo definido en la plantilla como su justificación, planteamientos legales, metodología, etc.

Para realizar esta investigación se realizan diversas investigaciones, ya sean artículos relacionados, y en base a estos contrastar diversas opiniones, puntos de vista, y argumentaciones de los autos, y para esto dar nuestro propio punto de vista, junto con su respectiva argumentación, y también dar soporte de dichos datos con tablas y/o gráficos, de igual manera, de emplear el aplicativo Hadoop el cual es de uso libre.

CAPÍTULO I: DESCRIPCIÓN DEL PROYECTO

Presentación del problema de investigación

Con la globalización del internet y su desarrollo a lo largo de los años, con esto presente, la investigación no solo tratará las redes sociales sino también tratará la masificación de la información, esta información no tiene como intención el ser socializada, sino más bien se tomara como punto de partida para el desarrollo del problema principal planteado , ya que este tema no tiene mucho que ver con el Big Data, pero en base a eso ya se han dado casos en los cuales la información sensible, filtraciones masivas, están relacionados con el mismo tema.

Con la investigación realizada con anterioridad se plantearán se plantea un problema general, el cual será la base de la investigación, como paso definiremos “¿Qué es la información sensible?”, la información sensible la podemos definir como aquellos datos personales de un individuo, ya sean credenciales bancarias, datos empresariales, entre otros, pero también en este mundo tan informatizado y que las personas suelen usar demasiado las redes sociales la información que se puede obtener de dichas redes sociales puede tratarse también como información sensible, por otro la información que se pueda encontrar en internet se puede considerar como información sensible, no solo temas globales como lo son, los números y contraseñas de las tarjetas de crédito de funcionarios gubernamentales, celebridades, entre otros, sino también dependiendo de la personas, sus credenciales de acceso a sus redes

sociales y/o correo personal el cual puede ser un tema supremamente delicado.

Con el concepto de los correos electrónicos y acotando las filtraciones masivas de correos relacionados a las redes sociales, y como estas pudieron tener puertas de acceso para la información haya sido extraída y posteriormente filtrada.

¿Es por estos y muchos más motivos que conocer como es el manejo real de la Big Data en ambientes que deben tener la mayor rigurosidad en su manipulación y cuidado de datos que se busca explorar nuestra incógnita, cómo se realiza el manejo de Big Data en redes sociales? Para esto se realizará una focalización en el área de marketing que vendría siendo una de las que más se beneficia del manejo de big data para esto tenemos que:

1. El marketing es un campo importante de la gestión de datos. Implica procesar y analizar grandes cantidades de datos para crear estrategias de marketing. Esto es esencial para que las empresas sigan siendo competitivas y para que los consumidores reciban los productos y servicios que desean. Gracias al desarrollo de la tecnología, el marketing se ha vuelto mucho más eficiente ya que utiliza software y hardware de análisis de datos.

En primer lugar, se recopilan y almacenan grandes cantidades de datos de muchas maneras, lo que los hace disponibles para su análisis con un detalle aparentemente infinito. Los datos se recopilan a través de diferentes medios, como redes sociales, mensajes de texto, historial de navegación y nombres de usuario. Además, se recopila a través de dispositivos como teléfonos móviles u ordenadores. Los datos recopilados luego se almacenan en bases de datos u hojas de cálculo. En ese punto, es fácil analizarlo y tomar decisiones con él.

Justificación

Dentro de lo primero que se busca plantear en esta investigación es ser capaz de generar ciertas pautas o guías que se usan en el estudio de big data en redes sociales, tomando como ejemplo, la analítica en el área o temática del marketing, es decir, lo que se busca es definir el proceso que se realiza para el estudio mencionado, lo que podría ser su descripción general para luego empezar a ahondar con su metodología de aplicación.

Para esto y como primer paso se busca encontrar diversas fuentes de información fiables que sirvan como base para la definición de los conceptos y pasos necesarios de este estudio, seguido se buscará filtrar la información obtenida para poder plasmar sólo la más importante en la generación de las pautas, una vez seleccionada esta información se apuntará a juntarla de forma coherente, como último se realizarán las validaciones pertinentes tanto a la información obtenida y sintetizada de las diferentes fuentes como al resultado

de la investigación, la definición de una metodología general hacia el estudio de big data en redes sociales el cual podrá servir como una guía.

Para poder realizar esta guía se utilizará uno de los programas más reconocidos y utilizados en este ámbito como viene a ser hadoop, se realizará una en esta guía una descripción no solo de los metodólogos generales de manipulación de big data sino también como se puede llegar a manejar a un nivel básico usando este programa, su forma de organizarlos, posibles añadidos con los que cuenta la herramienta tales como mapeo de resultados, manejo de lagos de datos, internet de las cosas, y etc.

Es decir, en resumen, se identificará que es y el manejo general del manejo de big data aplicándolo en la herramienta hadoop y como campo de ejemplo de estudio el "marketing".

Lo importante dentro de esto es saber cómo se le ha dado manejo a través de estos años al fenómeno creciente tecnológico el cual cuenta con información de todo tipo referente a una gran cantidad de personas pues, aunque se dé la ilusión que solo se almacenan contraseñas, correos e información básica de la persona en verdad esta información tiene una gran capacidad de ser usada tanto como forma de marketing como podría ser en base a conversaciones, audios, videos. Generar publicidad personalizada para la persona en base a lo que demostró generar interés, como forma de estafa frente a una persona amenazando con revelar datos privados al público o contactos por esta razón tener un conocimiento por lo menos básico de que es la big data cómo se

maneja en las redes sociales, como es usada la información registrada y en base a esto definir una metodología comúnmente usada explicada en la herramienta hadoop.

Con referencia a lo anterior se puede dar a conocer un poco como algunas empresas usan los datos que fueron recopilados con anterioridad y empieza a “mostrar” publicidad con base a la información que tienen de la persona, y con esto generar un ingreso monetario para dicha red social, y eso puede ser un problema con respecto a que tan invasivas pueden llegar las redes sociales hasta llegar al punto de usar distintos métodos para recopilar información para hacer marketing.

Objetivos

Objetivo general

Generar una investigación en la cual se pueda describir en base a la información recopilada que es, porque es importante y el manejo de big data en las redes sociales relacionados a su aplicación económica

Objetivos específicos

- Definir que es big data en general
- Generar una definición de cómo se maneja la información recopilada en las redes hacia un ámbito económico

- Introducir la definición y generar un acercamiento general de la herramienta conocida como apache hadoop

CAPÍTULO II: MARCO TEÓRICO

Antecedentes de la investigación

Primero se hará la definición general del tema a trabajar, es decir, la definición de que es big data, para que se usa esta misma, porque es importante su manejo y que metodologías se usan en la misma, después se sentarán las bases del uso de hadoop.

2.1 Antecedentes:

- Big Data Governance and Perspectives in Knowledge Management [1]
- La comunicación en la sociedad digital: reputación, big data-audiencia y medios, ante un público cambiante [2]
- La transformación del marketing sanitario [3]
- Beginning Apache Hadoop [4]
- Podcast: la importancia del marketing science y la inteligencia artificial con Avedis Boudakian [5]
- Marketing Digital en Colombia: Conoce su importancia [6]
- El concepto del Big Data en Colombia [7]
- Análisis del Big Data [8]

2.2 Bases teóricas

CAPÍTULO 1: Contexto de la situación

1.1. Definición

En los tiempos de hoy día el mundo está bajo un sistema masificado y globalizado por el internet, y teniendo en cuenta lo anterior no se tardó mucho en desarrollar las redes sociales como método de comunicación, ya sea de manera escrita y/o audiovisual. Donde, desde el principio se ha podido percatar la cantidad de usuarios que buscan hacer uso de estas herramientas de comunicación generando una gran demanda de recursos a las empresas propietarias de dichas redes sociales. o indirectamente recolecta estos mismos

1.2. ¿Qué es Big data?

Big data se conoce como el manejo y filtrado de la información masiva de un software con conexión a internet [9], ya sean redes sociales con sus imágenes, audios y mensajes o sitios de marketing que, aunque parezca que no son esta directamente relacionados son de los que más les interesa poder llevar un control y filtrado de la información para poder así con la ayuda de las apps de redes sociales generar anuncios de sus productos para llegar a un usuario que en base a la información que ha brindado está interesado en el producto en cuestión.

1.2.1. ¿Qué problemáticas trae la big data?

Aunque este manejo este dentro de una definición de marco legal, pues dentro de los términos y condiciones que contienen las paginas se especifica

que su información si bien no será divulgada ni compartida con agentes externos, esta sigue siendo usada por la misma compañía para poder ofrecer lo anterior mencionado, entrando en un territorio de legalidad y ética donde cualquier usuario puede llegar a sentir que su derecho a la privacidad está siendo violado

1.n. ¿Por qué son importantes estos conceptos?

Pues si bien dentro de la metodología del big data estos datos pueden sonar irrelevantes, si es necesario tenerlos a consideración en este contexto donde nos encontramos definiendo que es la big data como tal.

CAPÍTULO 2: Hadoop

2.1. ¿Qué es Hadoop?

Debemos entender que Apache Hadoop es un marco de software que brinda la capacidad de ejecutar aplicaciones distribuidas y escalables, a menudo utilizadas en el campo de los grandes datos. Como tal, permite que las aplicaciones aprovechen miles de nodos de procesamiento y almacenamiento y petabytes de datos.

Hadoop es una de las tecnologías más populares en el campo de las aplicaciones de big data. Se utiliza en muchas empresas como una plataforma central en sus lagos de datos, sobre los cuales se construyen casos de uso en torno al desarrollo y almacenamiento de datos.

Además, es una plataforma de desarrollo para aprovechar los datos de una organización, por ejemplo, a través de técnicas y modelos de aprendizaje automático.

Cada problema requiere una solución, por lo que acercarse a descubrir Apache Hadoop significa profundizar en los proyectos de Apache Software Foundation. Cada uno de ellos está desarrollado para proporcionar una funcionalidad clara y, como tal, cada uno tiene su propia comunidad de desarrolladores y un ciclo de lanzamiento independiente. El propósito de usar Hadoop es:

Para almacenar y procesar grandes cantidades de datos en una capa de almacenamiento escalable horizontalmente, nació Hadoop Distributed File System (HDFS). La tecnología funciona con hardware de bajo costo, sentando las bases para una escalabilidad eficiente desde la capa de almacenamiento. También se basa en Apache Hadoop HILLO, que proporciona administración de recursos y una arquitectura conectable para admitir múltiples métodos de acceso a datos para que los datos almacenados en Hadoop puedan manipularse con los niveles de servicio y rendimiento requeridos.

2.2. Uso de apache hadoop

Sistema de archivos distribuido Hadoop El sistema de archivos distribuido de Hadoop consiste en un sistema de archivos de alta disponibilidad para almacenar grandes cantidades de datos en clústeres de computadoras y es

responsable de mantener los datos dentro del marco. El clúster de Hadoop funciona básicamente de acuerdo con el principio maestro-esclavo. El marco del software consta de un nodo maestro, que está subordinado a varios nodos esclavos. Si esta señal no ocurre, el name node declara al esclavo "muerto" y asegura que, a pesar de este bloqueo, hay suficientes réplicas del bloque en cuestión en el clúster, utilizando las réplicas disponibles en otros nodos

Otro componente básico de Core Hadoop es el algoritmo MapReduce desarrollado por Google, que todavía se implementó como un motor separado en la primera versión de Hadoop. Las tareas centrales de esta máquina son la gestión de recursos y el control y seguimiento de los procesos informáticos. En la fase de mapa, el proceso de cálculo se divide en partes denominadas trabajos, y el rastreador de trabajos en el nodo maestro distribuye a cada sistema esclavo en el clúster. El siguiente diagrama muestra la arquitectura básica de Hadoop que funciona en la primera versión del marco, dividida en una capa MapReduce y una capa HDFS.

2.3. Como se aplica en el marketing

El manejo de datos orientado al marketing nos permite filtrar en base a temas que son recurrentes en este caso charlas de chat, grupos o detectado por voz lo que más se busca en general y a su vez lo que busca una persona en específico permitiendo de esta forma generar publicidad que ataca

directamente lo que la persona esta buscando para ofrecérselo y vender la información de la demanda actual a gran escala a las empresas encargadas.

Para esto y como se ya menciono anteriormente utilizando apache hadoop el cual tiene una gran cantidad de utilidades para su aplicación en los ámbitos que se quieran, en caso del marketing se puede utilizar una gran cantidad si no es que todas sus funcionalidades nombradas a continuación

HDFS

Sistema de archivos del disco duro Este es el componente del sistema de distribución de archivos Hadoop o algo similar que maneja el almacenamiento de datos. Cada archivo cargado en Hadoop se divide a su vez en fragmentos de 128 MB, que se pueden configurar para el almacenamiento en diferentes máquinas dentro del clúster de computadoras. A su vez, este componente de Hadoop tiene dos componentes principales: Nodo de nombre: este es el nodo maestro del que dependen todos los nodos de datos. Solo uno por grupo. Tu trabajo es saber dónde está cada pieza de información.

HIVE

Hive es un componente que facilita a los desarrolladores escribir funciones de MapReduce mediante consultas HQL. sqoop Muchas aplicaciones todavía usan bases de datos relacionales para almacenar información. En este caso, Sqoop es importante porque permite transferir esta información a HDFS. Pileta

FLUME

Flume es un servicio de código abierto para recopilar, agregar y mover de manera eficiente grandes volúmenes de datos de múltiples fuentes a HDFS. Este es uno de los componentes de Hadoop que puede recopilar datos en tiempo real y agruparlos por lotes.

Apache Kafka

Hay tantas aplicaciones que generan datos como aplicaciones que los consumen. Conectarlos es una tarea compleja. Aquí es donde entra Kafka. Kafka actúa como intermediario entre las aplicaciones que producen datos y las aplicaciones que los consumen. Por lo tanto, su función principal es distribuir datos entre ellos.

Oozie

Oozie es un sistema de programación que permite a los usuarios de Hadoop conectar fácilmente diferentes proyectos desarrollados en diferentes plataformas como MapReduce, Hive o Apache Pig.

Oozie permite programar el trabajo con anticipación y crear una serie de trabajos individuales para ejecutar secuencialmente o en paralelo para una tarea más grande.

Zookeeper

Zookeeper es un servicio centralizado, distribuido y de código abierto para mantener datos de configuración en Hadoop. Se utiliza para la asignación de nombres, la sincronización distribuida y la agrupación de servicios en un clúster.

Apache Spark

Spark es un marco alternativo a Hadoop basado en Scala que admite varias implementaciones de Java, Python y otros sistemas de programación de big data ampliamente utilizados.

Apache Spark ofrece a los expertos procesamiento por lotes de datos y puede procesarlos en tiempo real.

En resumen, estas herramientas nos ayudan a procesar de forma masiva la búsqueda de grandes grupos de personas sobre un tema en específico generando si un reporte o data de que es lo que está buscando la persona para poder ofrecer ese producto con más frecuencia en anuncios en básicamente todas las redes que puedan recompilar esta información por ejemplo buscar algo en youtube asegura que los próximos anuncios por ejemplo en google o Facebook estarán relacionados en este tema.

No solo eso, sino que también se podrá filtrar de mejor forma que tipo de cosas son las que está buscando una gran cantidad de gente así aunque no se tenga datos para generar una serie de publicidad personalizada si que se podrá generar un tema de interés que este en auge para ofrecerlo de forma más seguida a una gran cantidad de gente en general.

2.3 Bases legales.

Bases teóricas o fundamentos conceptuales

- El Estado Colombiano crea la Ley del Big Data en el documento Conpes 3920 que tiene como finalidad aumentar el aprovechamiento de datos,

mediante el desarrollo de las condiciones para su utilización como activos para generar valor social y económico [10]

- “Hace referencia al reconocimiento del valor social y económico de los datos, y a su comprensión como un activo e insumo para la producción de bienes y servicios. Este elemento es habilitador de los cambios institucionales requeridos para el consumo de datos, así como de los bienes, servicios y procesos generados o mejorados con estos” [11]
- “Hace referencia a las pautas de comportamiento que orientan a los actores del ecosistema hacia los fines socialmente deseables, de modo que se tomen las decisiones que maximicen el bienestar social en cada circunstancia particular.” [11]
- “Los datos son un insumo para la actividad investigativa, académica y productiva, para la toma de decisiones públicas y para la garantía de derechos. Son distintos a otros activos económicos en la medida en que su valor se materializa con el procesamiento y obtención de la información y el conocimiento intrínseco que se obtiene al combinarlos entre sí (European Data Protection Supervisor, 2014). Por tanto, el uso que se dé a los datos es el que, en últimas, materializa el valor potencial. Este activo reúne las características” [11]
- 44 Concepto definido como el deseo y habilidad de participar en la construcción de la sociedad usando datos (Data-Pop Alliance, 2015). Incluye el uso de soluciones de procesamiento y análisis de datos basadas en computadores u otros dispositivos, y la búsqueda, compartición, uso y creación de contenido usando TIC e Internet. [11]

Bases legales de la investigación

Constitución Política de Colombia	1991	Constitución Política de Colombia	
Ley	1266	2008	Por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones
Ley	1581	2012	Por la cual se dictan disposiciones generales para la protección de datos personales
Ley	1621	2013	Por medio de la cual se expiden normas para fortalecer el marco legal que permite a los organismos, que llevan a cabo actividades de inteligencia y contrainteligencia, cumplir con su misión constitucional y legal, y se dictan otras disposiciones
Ley	1712	2014	Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones.

Ley	1755	2015	Por medio de la cual se regula el Derecho Fundamental de Petición y se sustituye un título del Código de Procedimiento Administrativo y de lo Contencioso Administrativo
Ley	599	2000	Por la cual se expide el Código Penal
Ley	734	2002	Por la cual se expide el Código Disciplinario Único
Ley	1032	2006	Por la cual se modifican los artículos 257, 271, 272 y 306 del Código Penal (Ley 599 de 2000)
Ley	1273	2009	Por medio de la cual se modifica el Código Penal, se crea un nuevo bien jurídico tutelado - denominado "de la protección de la información y de los datos"- y se preservan integralmente los sistemas que utilicen las tecnologías de la información y las comunicaciones, entre otras disposiciones
Ley	1437	2011	Por la cual se expide el Código de Procedimiento Administrativo y de lo Contencioso Administrativo

Decisión	486	2000	Régimen Común sobre Propiedad Industrial
Decreto	3816	2003	Por el cual se crea la Comisión Intersectorial de Políticas y de Gestión de la Información para la Administración Pública COINFO
Decreto	575	2013	Por el cual se modifica la estructura de la Unidad Administrativa Especial de Gestión Pensional y Contribuciones Parafiscales de la Protección Social (UGPP) y se determinan las funciones de sus dependencias.
Decreto	602	2013	Por el cual se reglamenta el Sistema Único Nacional de Información y Rastreo (SUNIR), establecido por el Parágrafo 4° del artículo 227 de la Ley 1450 de 2011 y se dictan otras disposiciones
Decreto	235	2010	Por el cual se regula el intercambio de información entre entidades para el cumplimiento de funciones públicas
Decreto	2280	2010	Por el cual se modifica el artículo 3° del Decreto 235 de 2010
Decreto	1010	2000	Por el cual se establece la organización interna de la Registraduría Nacional del Estado Civil y se fijan las funciones de sus dependencias; se define la naturaleza jurídica del Fondo Social de Vivienda de la Registraduría Nacional del Estado Civil; y se dictan otras disposiciones
Decreto	857	2014	Por el cual se reglamenta la Ley Estatutaria 1621 del 17 de abril de 2013
Decreto	1377	2013	Reglamenta parcialmente la Ley 1581 de 2012
Decreto	886	2014	Registro Nacional de Bases de Datos
Decreto	1727	2009	Contenido mínimo de las historias crediticias
Decreto	2952	2010	Reglamenta artículos 12 y 13 la Ley 1581 de 2012
Decreto	805	2013	Registro de libros electrónicos de comercio
Decreto	2364	2012	Por medio del cual se reglamenta el artículo 7 de la Ley 527 de 1999, sobre la firma electrónica y se dictan otras disposiciones
Decreto	333	2014	Régimen de acreditación de las entidades de certificación
Decreto	489	2013	Por el cual se reglamenta parcialmente la Ley 1429 de 2010 (Por la cual se expide la Ley de Formalización y Generación de Empleo)
Decreto	504	1997	Por el cual se reglamenta el Registro Nacional de Turismo de que tratan los artículos 61 y 62 de la Ley 300 de 1996

Decreto	2183	2013	Por el cual se regula la elaboración y registro de las estadísticas relacionadas con el sector turístico y se dictan otras disposiciones
Decreto	103	2015	Disposiciones generales en materia de transparencia y del derecho de acceso a la información pública nacional
Decreto	2482	2010	Por el cual se dictan disposiciones relacionadas con la operación del Sistema de Información y Gestión del Empleo Público (SIGEP) y se deroga el Decreto 1145 de 2004.
Decreto	2482	2012	Por el cual se establecen los lineamientos generales para la integración de la planeación y la gestión
Decreto	415	2016	Por el cual se adiciona el Decreto Único Reglamentario del sector de la Función Pública, Decreto Numero 1083 de 2015, en lo relacionado con la definición de los lineamientos para el fortalecimiento institucional en materia de tecnologías de la información y las comunicaciones
Decreto	2052	2014	Por el cual se reglamenta la implementación del Sistema Único de Gestión e Información de la Actividad Litigiosa del Estado - eKOGUI
Decreto	1510	2013	Por el cual se reglamenta el sistema de compras y contratación pública
Decreto	1467	2012	Por el cual se reglamenta la Ley 1508 de 2012 (APP)
Decreto	1949	2012	Por el cual se reglamenta parcialmente la Ley 1530 de 2012 en materia presupuestal y se dictan otras disposiciones
Decreto	2844	2010	Por el cual se reglamentan normas orgánicas de presupuesto y del plan nacional de desarrollo
Decreto	1290	2014	Por el cual se reglamenta el Sistema Nacional de Evaluación de Gestión y Resultados - Sinergia-
Decreto	4816	2008	Por el cual se reglamenta el artículo 24 de la Ley 1176 de 2007 (SISBEN)
Decreto	3851	2006	Por el cual se organiza un sistema de aseguramiento de la calidad, almacenamiento y consulta de la información básica colombiana y se dictan otras disposiciones
Decreto	32	2013	Por el cual se crea la Comisión Nacional Digital y de Información Estatal
Decreto	489	2013	Por el cual se reglamenta parcialmente la Ley 1429 de 2010

Resolución	1424 7	2013	Por medio de la cual se modifica y aclara la Resolución número 3341 del 16 de abril de 2013.
Resolución	3564	2015	Por la cual se reglamentan los artículos 2.1.1.2.1.1, 2.1.1.2.1.11, 2.1.1.2.2.2, y el parágrafo 2 del artículo 2.1.1.3.1.1 del Decreto N° 1081 de 2015.
Resolución	2163	2016	Por la cual se reglamenta el Decreto 2297 de 2015 y se dictan otras disposiciones
Resolución	3313	2017	Por el cual se actualiza el Plan Vive Digital 2014-2018 contenido en el anexo de la resolución 828 del 11 de mayo de 2015 y se Deroga la Resolución 1564 de 2016
Circular	58	2009	Cumplimiento Decreto 1151 de 2008
Circular	52	2007	Requerimientos mínimos de seguridad y calidad en el manejo de información a través de medios y canales de distribución de productos y servicios para clientes y usuarios
Circular	2	2015	Adiciona capítulo segundo título V de la Circular Única de la SIC
Circular	5	2017	Adiciona capítulo tercero título IV de la Circular Única de la SIC
Sentencia	C 101 1	2008	Revisión de constitucionalidad del Proyecto de Ley Estatutaria No.27/06 Senado – 221/07 Cámara (Acum. 05/06 Senado)
Sentencia	C 274	2013	Revisión constitucional del Proyecto de Ley Estatutaria número 228 de 2012 Cámara, 156 de 2011 Senado, “por medio de la cual se crea la ley de transparencia y del derecho de acceso a la información pública nacional”
Sentencia	C 951	2014	Revisión de constitucionalidad del Proyecto de Ley número 65 de 2012 Senado y número 227 de 2013 Cámara “Por medio del cual se regula el derecho fundamental de petición y se sustituye un título del Código de Procedimiento Administrativo y de lo Contencioso Administrativo.”
Sentencia	T 729	2002	Acción de tutela instaurada por Carlos Antonio Ruiz Gómez contra el Departamento Administrativo de Catastro (Alcaldía Mayor de Bogotá) y la Superintendencia Nacional de Salud.

CAPÍTULO III: DISEÑO METODOLÓGICO

Tipo de investigación

- Propósito: Investigación básica.
- Lugar: Investigación documental.
- Alcance: Investigación descriptiva.

Definir:

- Cada uno de los tipos de investigación identificados anteriormente, conforme a los desarrollos del autor o teórico que el investigador tenga a bien considerar.

Argumentar:

- Por qué se identificó el tipo de investigación y de qué manera se relacionó con la propuesta presentada.
- Se identificó las características de la investigación como básica, documental y descriptiva por su naturaleza, pues se partió de una necesidad de consolidar un conocimiento sobre un tema más que investigar una solución a un problema real, dando así una característica de investigación básica mas no aplicaba pues no se busca satisfacer una problemática.
- Así mismo como ya se demostró en el documento esta no intenta mostrar un punto de vista o generar una comparativa, se logró resaltar hechos reales y aplicables basados en la investigación del tema en cuestión la big data relacionada a su delimitación y facilitador es decir el marketing y la descripción de una herramienta de uso común en el ámbito.

- Con fin de demostrar la aplicación de los datos obtenidos mediante el manejo de big data en las redes sociales se tomó como ejemplo su aplicación en el área económica

Población

La población se orientó para que este documento sea legible para cualquier persona interesada por el tema, que pueda tener un punto de partida hacia la investigación más profunda de este mismo tema, en cuanto a la población relacionada en el campo de investigación se intentó focalizarla en fuentes confiables que hablen de lo anteriormente mencionado en el proyecto de investigación

Técnicas e instrumentos de recolección de datos

Se utilizó el apoyo de un tutor para poder encaminar el tema de investigación por el cual seguir, así mismo todo el documento fue realizado en base a fuentes encontradas en internet, no solo paginas sino artículos, libros, archivos pdf y definiciones gubernamentales del tema en cuestión

CAPÍTULO IV: RESULTADOS DE LA INVESTIGACIÓN

Resultados del objetivo específico no. 1

A lo largo del desarrollo del documento se fue recopilando y explicando específicamente tanto el manejo de big data en general, su definición uso, surgimiento y características como su específico orientado a redes sociales la cual también fue definida dentro del contenido tanto una explicación general, como se dio, para que se usa, posibles aplicaciones, afectaciones e intervenciones en negocios con su manejo dado en empresas.

Resultados del objetivo específico no. 2

Gracias al apoyo de la herramienta ya varias veces mencionadas hadoop se logró ver como esta facilita los procesos de obtención, clasificación y filtrado de los datos, para su posterior uso, un ejemplo de esto fue el ya mencionado Map Reduce que nos permite dividir un procesamiento masivo en múltiples partes más pequeñas y su descripción de como se hace uso en el post o resultado de este procesado.

Resultados del objetivo específico no. 3

Como ya se mencionó se exploró lo que viene a ser la herramienta hadoop, mas sin embargo no se le dio un uso práctico como tal pues la data necesaria y el tiempo de aprendizaje de la misma herramienta es algo bastante amplio, necesitando a su vez un ambiente de sistemas Linux por lo que cualquier proyección que se hubiera propuesto quedaría corta o incompleta, más a pesar de esto se compilo las suficientes descripciones y capacidades como para dar solución a la parte de definir el alcance de esta herramienta junto a como se realiza el manejo en un tema como el marketing.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

Como conclusión a lo largo de todo este desarrollo se logró concentrar la información pertinente al tema en cuestión la big data, su definición orientada a las redes sociales y como esta información es usada, junto a métodos de sintetizarla definidos en la parte de la herramienta apache hadoop, como se logra observar en los capítulos definidos en su respectiva parte se definió con varias fuentes el uso de big data en redes sociales apoyados en el área de marketing con las facilidades que ofrecer la herramienta hadoop.

Se logró conseguir satisfactoriamente lo propuesto, las definiciones del tema en cuestión con una orientación aplicable sintetizado de múltiples fuentes de información generando este documento el cual da una noción bastante clara para cualquier persona que se disponga a leerlo.

Bibliografía

- [1 S. Kruger Strydom y M. Strydom, Big Data Governance and Perspectives in Knowledge Management, IGI Global, 2019.
- [2 CEU Universidad San Pablo, La comunicación en la sociedad digital: reputación, big data-audiencia y medios, ante un público cambiante, ESIC, 2020.
- [3 J. M. Martínez García y L. Pérez Campillo, La transformación del marketing sanitario, Madrid: ESIC, 2021.
- [4 M. Road y Chetpet, Beginning Apache Hadoop, Notion Press, 2017.
-]

[5 RevistaPyM, «Podcast: la importancia del marketing science y la inteligencia artificial con Avedis Boudakian,» PyM, 2022.

[6 S. A. Digital, «simbolointeractivo,» 2022. [En línea]. Available: <https://www.simbolointeractivo.com/marketing-digital-en-colombia/>.

[7 S. M. Jaramillo, «asuntoslegales,» 2022 11 18. [En línea]. Available: <https://www.asuntoslegales.com.co/consultorio/el-concepto-del-big-data-en-colombia-3068179#:~:text=El%20uso%20de%20Big%20Data,fortalecer%20su%20relaci%C3%B3n%20con%20stakeholders..>

[8 Ibm, «IBM,» 2022. [En línea]. Available: <https://www.ibm.com/com-es/analytics/hadoop/big-data-analytics>.

[9 J. Lopez y E. Ruiz, Libro de big data, Bogotá, D.C.: Planeta, 2021.

[10] L. A. S. HERNANDEZ, «<https://bibliotecadigital.univalle.edu.co/>,» 2019. [En línea]. Available: <https://bibliotecadigital.univalle.edu.co/bitstream/handle/10893/14570/3340-0582743-E.pdf?sequence=1&isAllowed=y#:~:text=El%20Estado%20Colombiano%20crea%20la,generar%20valor%20social%20y%20econ%C3%B3mico..>

[11] C. N. D. P. E. Y. SOCIAL, «POLÍTICA NACIONAL DE EXPLOTACIÓN DE DATOS,» 2018. [En línea]. Available: <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3920.pdf>.

