

**Creación de Procedimiento de Big Data para el Ministerio de Justicia y
del Derecho.**

Autor

Andrés Felipe Méndez Ribero

Laura Marcela Patiño Gutiérrez

Docente Tutor ECACEN UNAD

Universidad nacional abierta y a distancia UNAD

Escuela de Ciencias Administrativas, Contables Económicos y de

Negocios (ECACEN) especialización gestión de proyectos

DEDICATORIA

Le Dedico este trabajo de Grado a mis padres que fueron un pilar importante para mí desarrollo académico, pues sin ellos no sería el profesional que soy ahora. También a mi esposa y mi hija que son el motor para seguir adelante y cumplir las metas que me proponga y final mente a mi familia en general y a mi tutora que me ayudo en este camino para culminar con este proyecto.

Tabla de contenido

Resumen analítico especializado RAE	6
Resumen	9
Abstract.....	12
Introducción	13
Objetivo general	15
Capítulo 1_¿Que es el Big Data y para que sirve?	16
Capítulo 2_Antecedentes Big Data en Colombia.....	21
Capitulo 3_Como la Big Data apoya los procesos internos del Ministerio de Justicia y del Derecho ..	28
Capítulo 4_Componentes para la plataforma Big Data del Ministerio de Justicia y del Derecho	38
Capitulo 5_Estrategia para aprovechar la Big Data en el Ministerio de Justicia y el Derecho	48
Capitulo 6_Seguridad en la explotación de los datos.	53
Capitulo 7_Cultura de datos.....	57
Capitulo 8_Balanced Scorecard Big Data Ministerio de Justicia y del Derecho	63
Conclusiones	66
Bibliografía	69

Índice De Figuras

Figura N°1 Crecimiento de datos en el Mundo Dirección Nacional de Planeación, IBM OCDE 2015	17
Figura N° 2 Marco Ético - Jurídico Estacional (Big Data)	18
Figura N° 3 Barreras para poder Intercambiar datos en el Sector Publico, Encuesta DNP (2017), Elaboración propia. Universo: 209 entidades (orden nacional). Muestra: 150 entidades	27
Figura N° 4 Pagina web Ministerio de Justicia y del Derecho/organigrama	29
Figura No. 5 Flujograma de Procesos Big Data, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co	31
Figura No. 6 Flujograma de Procesos Big Data Viceministerio Promoción a la Justicia, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co	33
Figura No. 7 Flujograma de Procesos Big Data Viceministerio Política Criminal y Justicia Restaurativa, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co	35
Figura No. 8 Flujograma de Procesos Big Data Oficinas de apoyo, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co	36
Figura No. 9 Flujograma de Procesos Big Data Oficinas de apoyo, Ministerio de Justicia y del Derecho Organigrama.....	37
Figura No. 10 Estructura de una Herramienta ETL, Data Prix Herramientas ETL. ¿Que son, para que valen? Productos más conocidos. ETL´s Open Source.	39
Figura No. 11 Tratamiento de Datos Fundamentos de Apache Hadoop y MapReduce https://geekytheory.com/fundamentos-de-apache-hadoop-y-	43
Figura No. 12 Base de datos distribuida Elasticsearch, Importación y análisis de contrataciones en formato Estándar de Datos de Contrataciones Abiertas (OCDS) con herramientas Elastic,	45
Figura No. 13 Elasticsearch Cluster, Importación y análisis de contrataciones en formato Estándar de Datos de Contrataciones Abiertas (OCDS) con herramientas Elastic,	46
Figura No. 14 Flujograma de Procesos Organizacional Big Data, MyABCM ¿Cómo puede su empresa ser innovada por la cultura basada en datos?	58
Figura No. 15 Balance ScoreCard Big Data MJD.....	63

ÍNDICE DE TABLAS

Tabla N°1 Mapa Estratégico Cultura de Datos	49
---	----

RESUMEN ANALITICO ESPECIALIZADO RAE	
1. INFORMACIÓN GENERAL	
Tipo de Documento	Monografía Trabajo opción de Grado
Tema	Diseño y creación de los lineamientos que debe tener el Ministerio de Justicia y del Derecho para implementar la Big Data en sus procesos Internos
Título	Creación de procedimiento de big data para el ministerio de justicia y del derecho
Autor	Andres Felipe Mendez Ribero
Director	Laura Marcela Patiño Gutiérrez
Identificación	Cedula de Ciudadanía No. 1098678279
Programa Académico	Especialización Gestion de Proyectos
Escuela	Escuela de Ciencias Administrativas, Contables, Económicas y Negocios-ECACEN
Zona	Centro
Ciudad	Bogotá D.C.
Año	2020
Resumen	La importancia de desarrollar la creación de una Big Data para el Ministerio de Justicia radica en apoyar y aprovechar las nuevas tendencias en la gestión de la información, lo cual permitirá realizar análisis más detallados a los distintos casos que se puedan presentar y que no puedan ser resueltos mediante bases de datos convencionales y con esto se fortalecerá la temática para posteriores vinculaciones en solicitudes que se tengan a nivel interno o externo. Por esto la intervención de esta política al Ministerio de Justicia es especialmente necesaria en situaciones donde las presiones competitivas son bajas o inexistentes, como ocurre en el sector público. En otras palabras, es necesario que las condiciones para la explotación de datos sean impulsadas mediante la intervención pública, corrigiendo las fallas de gobierno que impiden el surgimiento de elementos que faciliten el acceso a la información de manera fácil y rápida. Estos datos se deben recopilar de manera rutinaria y masiva, que por su naturaleza no es creado por el mercado si no por las necesidades y requerimientos que tenga la entidad lo cual se convierten en datos públicos que ayuden a un mejor entendimiento de los procesos judiciales de la nación.
Palabras Claves	Big Data, Ministerio de Justicia y del Derecho, Estrategia Tecnológica, Información, Documentación, Información Pública, Acceso a información, procesos Internos, Cultura de Datos, Seguridad en Datos.
Contenido	<ul style="list-style-type: none"> • Objetivo general y específico • Capítulo 1 ¿qué es el Big Data y para qué sirve? • Capítulo 2 antecedentes Big Data en Colombia • Capítulo 3 como la big data apoya los procesos internos del Ministerio de Justicia y del Derecho

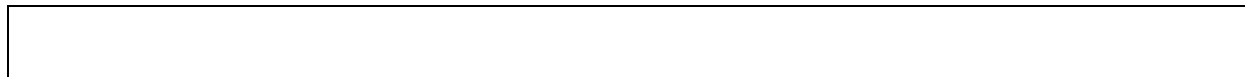
	<ul style="list-style-type: none"> • Capítulo 4 componentes para la plataforma Big Data del Ministerio de Justicia y del Derecho • Capítulo 5 estrategia para aprovechar la Big Data en el Ministerio de Justicia y el Derecho • Capítulo 6 seguridad en la explotación de los datos • Capítulo 7 cultura de datos • Capítulo 8 balanced ScoreCard big data Ministerio de Justicia y del Derecho • Conclusiones • Bibliografía
Objetivo General	Formular una estrategia para mejorar la gestión de información en el Ministerio de Justicia y del Derecho mediante el concepto de Big Data.
Objetivos Específicos	<ul style="list-style-type: none"> • Diseñar una estrategia de Masificación de la disponibilidad de datos para la entidad que sean digitales accesibles, usables y de calidad. • Identificar tecnologías, herramientas de software y requerimientos de hardware necesarios para la implementación de la estrategia de gestión de Big Data en el MJD¹ • Diseñar una estrategia de cultura de datos que facilite la gestión del Big Data en el MJD
Metodología	Desde la vigencia 2018, se comenzó la idea de crear una base para poder implementar la Big Data en los procesos del Ministerio de Justicia y del Derecho. Con este orden de ideas se hizo un acercamiento con las áreas de tecnologías de la entidad para identificar cual podría ser el camino para llevar a cabo el desarrollo del proyecto. También se apoyó con el Ministerio de Tecnológicas de la Información y Comunicaciones dando recomendaciones en temas de seguridad de los datos y que se tuviera en cuenta el documento conpes 3920 donde se menciona la política de explotación de Datos para las entidades de orden Nacional.
Principal(es) referentes teóricos	El principal Referente Teórico es el Documento Conpes 3920 dirigido entre el Departamento Nacional de Planeación y el Ministerio de la Información y comunicaciones. Estas entidades bajo este documento lideran la idea de implementar una política de explotación de datos para generar un mejor aprovechamiento de datos creando al mismo tiempo un valor social y económico para las entidades nacionales.
Resultados	Se espera que el Ministerio de Justicia y el derecho tenga un procedimiento adecuado para poder implementar la Big Data en sus procesos internos y crear un valor agregado a la entidad. Esto con el fin de brindar información efectiva, veras y ágil a los requerimientos que se tengan por parte de los ciudadanos o la Nación.
Conclusiones	La Big data ayudara al Ministerio de Justicia y del Derecho a crear un mecanismo donde pueda agilizar sus procesos internos, guardando información física en respaldos digitales para así poder crear una red de información generosa y que cualquier entidad estatal necesite y pueda solicitarla de manera

más rápida y segura, enfocada en temas que tengan que ver con el sector justicia. La Big Data ayudara a la entidad a dar respuestas más rápidas agilizando los canales de recepción y envío de información.

Esta Data digital ayudara a tener un respaldo más eficiente que el fisco ya que no es necesario desplazarse donde está este material el cual está guardado en archivos físicos, si no que puedan ser consultados por personal autorizado. En la red se busca que la información que llegue ayude a la entidad a generar un pronóstico de la situación actual del país la cual genere una política Jurídica a los casos más buscados o consultados por los ciudadanos o por las empresas, este es el fin de la Big Data en el Ministerio de Justicia y del Derecho.

FUENTES BIBLIOGRAFICAS

- ACENS. (s.f.). *Bases de datos NoSQL, Qué son y tipos que nos podemos encontrar*. Obtenido de Acens: <https://www.acens.com/wp-content/images/2014/02/bbdd-nosql-wp-acens.pdf>
- Arias, D. (2018). *Colombia es el primer país en explotar Big Data en Latinoamérica*. Obtenido de Enter.co: <https://www.enter.co/especiales/empresas/colombia-primer-exploracion-de-big-data/>
- Barranco Frago, R. (2012). *¿Que es Big Data?* Obtenido de IBM: <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html>
- Blanco, T. P., & Sánchez Herrera, J. (2014). *Nuevas tendencias en comunicación estratégica*. Madrid: ESIC EDITORIAL.
- D'Adamo, L. (2018). *Elasticsearch: qué es, cómo funciona y características*. Obtenido de Consulthink: <https://www.consulthink.it/es/elasticsearch-que-es-como-funciona-y-caracteristicas/>
- Departamento Nacional de Planeación. (2018). *Política Nacional de Explotación de Datos (Big Data)*. Obtenido de <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%20B3micos/3920.pdf>
- Departamento Nacional de Planeación. (s.f.). *Bases del Plan Nacional de Desarrollo 2014-2018*. Obtenido de <https://colaboracion.dnp.gov.co/cdt/prensa/bases%20plan%20nacional%20de%20desarrollo%202014-2018.pdf>
- DíazGranados Ordóñez, G. F. (2016). *La comercialización de la Big Data*. Obtenido de Ciencias Jurídicas Javeriana: <https://cienciasjuridicas.javeriana.edu.co/documents/3722972/7912168/7-%20LA+COMERCIALIZACION%20de%20Big+Data.pdf/17c8d327-d25b-474c-a9c1-1797114e3c0b>
- Ministerio de Justicia y del Derecho. (s.f.). *Estructura Organizacional de la Entidad*. Obtenido de <https://www.minjusticia.gov.co/Ministerio/Estructura-organizacional-de-la-entidad>
- Ministerio de Tecnologías de la Información y las Comunicaciones. (s.f.). *Sistemas de Gestión de la Seguridad de la Información (SGSI)*. Obtenido de <https://www.mintic.gov.co/gestion-ti/Seguridad-TI/Modelo-de-Seguridad/5482: Sistemas-%20de-Gesti%C3%B3n-de-la-Seguridad-de-la-Informaci%C3%B3n-SGSI>
- MyABCM. (s.f.). *¿Cómo puede su empresa ser innovada por la cultura basada en datos?* Obtenido de MyABCM: <https://www.myabcm.com/es/blog-post/cultura-basada-en-datos/>
- Valero, M. (2016). *¿Que es una ETL?* Obtenido de Biggeek: <https://blog.bi-geek.com/que-es-una-etl/>
- Vega, R., Szyszlican, M., & Borregón, E. M. (s.f.). *Manual Kibana OCDS*. Obtenido de https://manualekibanaocds.readthedocs.io/es/latest/C2/Seccion1/1_ElasticSearch.html
- Villarrubia, C. (2018). *Colombia define una política pública para big data*. Obtenido de Datacenterdynamics: <https://www.datacenterdynamics.com/es/noticias/colombia-define-una-politica-publica-para-big-data/>



RESUMEN

El proceso de datificación se debe implementar en el Ministerio de Justicia y

el Derecho ya que se quiere transformar la documentación en general de la entidad en datos procesables y cuantificables, dando lugar a la generación de nuevos bienes, servicios y procesos y así aportar la diversificación y sofisticación que redonda en el crecimiento de la era digital que estamos viviendo.

La importancia de desarrollar la creación de una Big Data para el Ministerio de Justicia radica en apoyar y aprovechar las nuevas tendencias en la gestión de la información, lo cual permitirá realizar análisis más detallados a los distintos casos que se puedan presentar y que no puedan ser resueltos mediante bases de datos convencionales y con esto se fortalecerá la temática para posteriores vinculaciones en solicitudes que se tengan a nivel interno o externo. Por esto la intervención de esta política al Ministerio de Justicia es especialmente necesaria en situaciones donde las presiones competitivas son bajas o inexistentes, como ocurre en el sector público. En otras palabras, es necesario que las condiciones para la explotación de datos sean impulsadas mediante la intervención pública, corrigiendo las fallas de gobierno que impiden el surgimiento de elementos que faciliten el acceso a la información de manera fácil y rápida. Estos datos se deben recopilar de manera rutinaria y masiva, que por su naturaleza no es creado por el mercado si no por las necesidades y requerimientos que tenga la entidad lo cual se convierten en datos públicos que ayuden a un mejor entendimiento de los procesos judiciales de la nación.

Por otro lado, las intervenciones en materia de protección de datos personales se enfrentan al desafío de garantizar los derechos asociados, en un contexto de aumento exponencial en las formas de generación y recolección de estos datos en formato digital,

unido a la sofisticación de las capacidades para aprovecharlos lo cual mejorará la transparencia e incrementará el acceso a la información pública, así como la eficiencia y eficacia en la intervención pública a través de la toma de decisiones mejor informadas.

ABSTRACT

When we talk about Big Data, we refer to data sets or combinations whose size (volume), complexity (variability) and speed of growth (speed) make it difficult to capture, manage, process or analyze using conventional technologies and tools, such as data bases. Relational data and conventional statistics or visualization packages, within the time necessary for them to be useful and thus mitigate any request that is required and act in an agile manner to weigh any requirement. The purpose of this monograph is to publicize a tool that helps the Ministry of Justice and Law in streamlining its internal and external processes in judicial matters and that generates added value that can serve as a model for other government entities. The previous interventions that are desired in electronic management, administrative efficiency and document management issues give rise to the initial conditions for data exploitation. With the document Conpes 3920 where it determines the first policy of data exploitation in the country, it opened the doors for the entity to enter to use the virtual media and data collection of users maximizing the mechanisms that they have, and this will help to the management of both physical and virtual information. We can observe that the current situation regarding the low availability of public data is divided into three situations: low digitalization, low levels of open data and low interoperability, also reflecting the lack of articulation of defined policies and insufficient intervention in the manipulation of digital data, since there are no specific tools and mechanisms that allow the creation of a database that is easily accessible to officials. The projection that is had is that in 4 years the Ministry of Justice has a Big Data project that generates social and economic value to the nation.

INTRODUCCIÓN

Cuando hablamos de Big Data nos referimos a conjuntos de datos o combinaciones de cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles y así mitigar cualquier solicitud que se requiera y actué de manera ágil para sopesar cualquier requerimiento. El propósito de esta monografía es dar a conocer una herramienta que ayude al Ministerio de Justicia y del Derecho en agilizar sus procesos tanto internos como externos en temas judiciales y que le generen un valor agregado que pueda servir de modelo para las demás entidades gubernamentales. Las intervenciones previas que se desean en temas de manejo electrónico, eficiencia administrativa y gestión documental dan lugar a las condiciones iniciales para la explotación de datos. Con el documento Conpes 3920 donde determina la primera política de explotación de datos en el país abrió las puertas para que la entidad entre a utilizar los medios virtuales y recopilación de datos de los usuarios maximizando los mecanismos que necesidades que se tengan, ya con esto ayudara al manejo de información tanto física como virtual.

Podemos observar que la situación actual en cuanto a la baja disponibilidad de datos públicos se divide en tres situaciones: baja digitalización, bajos niveles de datos abiertos y baja interoperabilidad, también reflejando la ausencia de articulación de las políticas definidas e intervención insuficiente en la manipulación

de datos digitales, ya que no se dispone de herramientas y mecanismos concretos que permitan la creación de una base de datos que sea de fácil acceso para los funcionarios. La proyección que se tiene es que en 4 años el Ministerio de Justicia tenga un proyecto de Big Data que generará valor social y económico a la nación.

OBJETIVO GENERAL

Formular una estrategia para mejorar la gestión de información en el Ministerio de Justicia y del Derecho mediante el concepto de Big Data.

OBJETIVO ESPECIFICO

1. Diseñar una estrategia de Masificación de la disponibilidad de datos para la entidad que sean digitales accesibles, usables y de calidad.
2. Identificar tecnologías, herramientas de software y requerimientos de hardware necesarios para la implementación de la estrategia de gestión de Big Data en el Ministerio de Justicia y del Derecho.
3. Diseñar una estrategia de cultura de datos que facilite la gestión del Big Data en el MJD

CAPÍTULO 1

¿QUE ES EL BIG DATA Y PARA QUE SIRVE?

Para empezar la datificación son la representación primaria de variables almacenadas, transferibles, que pueden visualizar, controlar y comprendidas. Los datos que se perciben por los sentidos humanos se denominan analógicas, pero si las codificamos en un sentido binario podemos transformarlas en datos digitales. Este proceso de recepción de datos representa una oportunidad para genera una mayor información y conocimiento que se traduce en valor social y económico para un país

El crecimiento de datos en el mundo tiene una tendencia al alza registrada por IBM, el cual estimo que el volumen de datos diarios es alrededor de 2,5 trillones de bytes, en términos más entendibles la cantidad de datos que se producen equivalen más o menos a la capacidad de almacenamiento de 3.3 millones de CD'S por día. Uno de los periodos donde más alto se detecta a nivel mundial el uso de datos fue en el 2013 comprendiendo en un 81% y que desde ese año hasta el 2018 esa cifra ha crecido casi en 11 veces (Blanco & Sánchez Herrera, 2014)¹.

Según IBM para el año 2020, se habrá generado en el mundo la impresionante cantidad de 40 Zeta bytes o 40 millones de gigabytes de contenido digital. Hoy en día la información se acumula de forma sorprendente teniendo un punto de partida en las actualizaciones en redes sociales, hasta el punto de obtener

¹ Teresa Pintado Blanco – Joaquín Sánchez Herrera, Nuevas Tendencias en comunicación estratégica 4º Edición, (ESIC – Business Marketing School) (p.374).

información de las compras realizadas por red. Se puede afirmar que el 90% de los datos actuales han crecido en los últimos dos o tres años.

Un ejemplo importante es que desde el 2013 se sube a Twitter una plataforma para escribir mensajes cortos de opinión más de 300 millones de mensajes al día.

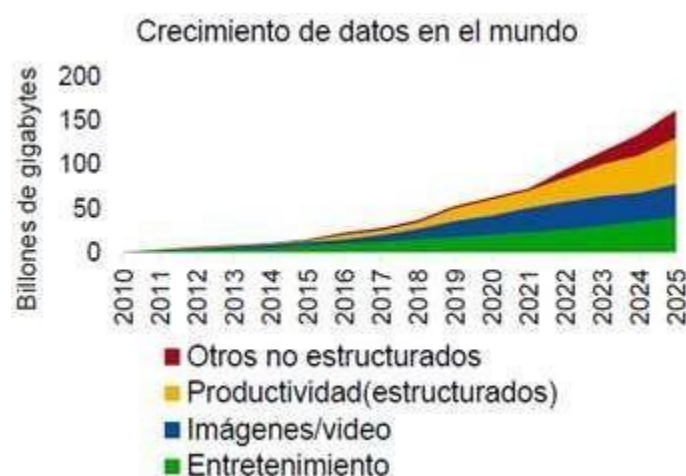


Figura N°1 Crecimiento de datos en el Mundo Dirección Nacional de Planeación, IBM OCDE 2015

En la política de explotación de datos Colombia es el primer país en América Latina y el octavo en el mundo en adoptar una política nacional de explotación de datos, estas políticas buscan proveer condiciones habilitantes para lograr la generación de valor a la vida social y económica del país.

En el Conpes 3920 (Departamento Nacional de Planeación, 2018), que es el documento que expresa la política nacional de explotación de datos (BIG DATA) busca el aprovechamiento de los mismos para aprovecharlos y generar un valor social y económico para las entidades del estado. El marco ético, Jurídico e institucional en:

Datos Digitales = Cultura de datos = Valor Social y Económico (Producto)

Capital Humano para la explotación de Datos

Figura N° 2 Marco Ético - Jurídico Estacional (Big Data)

El concepto de Big Data aplica para toda aquella información que no puede ser procesada utilizando procesos o herramientas tradicionales, sin embargo, se tiene que tener claro que Big Data no se refiere a la cantidad en específico, ya que es usualmente utilizado cuando se habla en términos de peta bytes y exabytes de datos, por lo cual se dará a conocer la cantidad de bytes que establecerá cual podríamos establecer para ser procesada y analizada utilizando Big Data (Barranco Fragoso, 2012) :

Gigabyte = 10^9 =

1.000.000.000 Terabyte =

10^{12} = 1.000.000.000.000

Peta byte =

1.000.000.000.000.000

Exabyte =

1.000.000.000.000.000.000

Esta cantidad de espacio para guardad información está representada en una gran variedad de datos que pueden ser representados en la entidad de diferentes maneras, como por ejemplo en datos de audio, video, documentos ofimáticos los cuales pueden medir y comunicar el estado de las dependencias que conforman el Ministerio de Justicia y del Derecho. Esto es una de las características que podemos aprovechar para crear la Big Data de la entidad.

Debemos tener en cuenta que las bases de datos convencionales son una parte muy importante para una solución analítica. De hecho, es muy vital cuando se usa en conjunto con la plataforma Big Data. Una de las analogías es pensar en nuestras dos manos, cada una ofrece fortalezas individuales para cada tarea que realicemos, podemos relacionarla por ejemplo con un encuentro de ultimate el cual está basado en capturar un disco y llevarlo a la cancha contraria para anotar un punto en este sentido necesitamos de nuestras manos para lanzar el disco y en otro caso para atraparlo, por lo cual nos lleva a que se necesita de las bases de datos y la Big Data para crear un control más estricto y claro de las cosas.

TIPOS DE DATOS

Existen tres grandes bloques o tipos de datos donde podemos realizar la clasificación de los ítems necesarios:

Datos no estructurados: Los datos en el formato original que no fueron recopilados, no se pueden almacenar en tablas y no se puede simplificar. (En esta categoría entrarían datos como videos, audios, multimedia, PDF, correos electrónicos).

Datos Estructurados: Los datos con formato definidos, que contengan fechas, números o cadenas de caracteres. Estos se almacenan a este tipo de datos en tablas, como por ejemplo Hojas de cálculo y datos transaccionales.

Datos Semiestructurados: Datos combinados que no se limitan en campos determinados, estos datos están catalogados con información irregular y son datos que poseen otros datos que se describen entre ellos, un ejemplo son los datos de programación (páginas web).

CAPÍTULO 2

ANTECEDENTES BIG DATA EN COLOMBIA

La instalación de un ambiente que utiliza la herramienta Big Data puede contener temáticas ambiguas, debido a la plataforma o uso que se requiera, debido a la cantidad de aplicaciones que se tienen y el manejo diferente en sus usos. Durante los años 90 y principios y mediados del 2000 se quería implementar el uso de las TIC'S en maximizar la eficiencia en el gesto de las entidades públicas de la nación. Sin embargo, desde 2017 se reconoce en Colombia que las TIC'S aumentan en gran medida la eficiencia y eficacia de la labor de las entidades.

Pero para poder lograr entender cuál ha sido el camino largo y tedioso de todo este asunto debemos entender que en la década de los 1997 en colombina ya se quería reconocer a las Tecnologías de la información y la comunicación (TIC'S) como parte fundamental en la implementación y aprovechamiento del desarrollo y evolución digital de ese entonces para poder tener un punto de inflexión tanto en el manejo grande de datos que se tenía y guardarlos en un lugar de rápida consulta dando así un valor agregado como la eficiencia de las entidades públicas y dar un servicio de calidad a los usuarios.

Ya en la década del 2000, se cambia el paradigma de brindar un servicio de calidad al ciudadano volviendo realidad, con el concepto de la nueva era digital que se venía venir. Las temáticas específicas de las estrategias de Gobierno Electrónico que era

una política que el gobierno quería implementar en esta nueva era de desarrollo, lo cual buscaba principalmente en el aprovechamiento de estos datos convirtiéndolos en no solamente copias de estos mismos sino respaldos de documentos digitalizados y ubicarlos en lugares donde podrían reposar seguros. Esto desencadenó un punto de partida para el País en tema de manejo de datos masivos digitalizados.

En el 2003, mediante el Conpes 3248² (Renovación de la administración Pública, nos da a conocer los puntos de interés para adecuar la administración pública y fortalecer su gestión, sin embargo, acá se debe tener en cuenta que para esta transformación pública se debe incluir el Gobierno Electrónico lo cual generó un cambio sustancial en la baja calidad y desactualización del manejo de información, así como la incompatibilidad tecnológica que existía generando un fortalecimiento de esta área.

Como resultado se implementaron la creación del sistema Nacional de información oficial, lo cual generó las bases del Gobierno Electrónico, esta política fue establecida en la Ley 790 de 2000 y adaptada en la implementación a las entidades públicas mediante el Decreto 1151 de 2008 juntando todas estas características para tener un punto de partida más concreto.

Ya para el año 2010 y mediante el conpes 3650³ donde se toca la implementación

² El conpes 3248, fue una política establecida en el primer periodo presidencia del Presidente Álvaro Uribe Vélez, la finalidad de esta política de Gobierno Electrónico fue definir los instrumentos adecuados para el manejo de la información en el sector público.

³ En el conpes 3650, establece los lineamientos generales de la estrategia de gobierno en línea en Colombia, el cual establece los principios en: Tener un Gobierno Centrado en el Ciudadano, Acceso equitativo, Protección de la información al ciudadano.

estratégica de la estrategia de gobierno en línea es donde se resalta la responsabilidad de cumplir con los lineamientos del Decreto 1151 de 2008. Se quiso expandir la idea de la mejora en el servicio ofrecido en línea a los ciudadanos. De acuerdo a lo anterior y mediante el Decreto 2695 de 2012, se actualizaron los componentes que tenía planificado el proyecto Gobierno en línea, integro la interoperabilidad entre entidad usuario, ventanillas virtuales que mejoraban las búsquedas que se requieran, todo esto para generar datos Virtuales que ayudaran a las entidades Públicas en temas tanto de atención al cliente como en tener un mejor manejo de datos y así poder tener todo a la mano generando un valor agregado tanto para rendir de cuentas como para solicitudes de externos.

Ya para el año 2018 se suscribe el Conpes 3920 donde se expresa claramente la explotación de BIG DATA en las entidades. La presente política tiene por objetivo aumentar el aprovechamiento de datos, mediante el desarrollo de las condiciones para que sean gestionados como activos para generar valor social y económico. En lo que se refiere a las actividades de las entidades públicas, esta generación de valor es entendida como la provisión de bienes públicos para brindar respuestas efectivas y útiles frente a las necesidades sociales.

En el artículo 15 de la constitución política de Colombia, establece el decreto del Habeas Data, que menciona lo siguiente:

De igual modo, tienen derecho a conocer, actualizar y ratificar las informaciones que se hayan recogido sobre ellas en bancos de datos y en archivos de entidades públicas y privada. En la recolección, tratamiento y circulación de datos se

respetarán la libertad y demás garantías consagradas en la constitución.

Y dando desarrollo al artículo 15 estando a la mano con los Estándares Internacionales sobre protección de Datos personales, la Ley 1581 de 2012, regula la protección de datos personales. Entre sus disposiciones se encuentran el tratamiento de datos personales, categorías especiales para los datos, derechos y condiciones para el tratamiento de datos, deberes y responsables encargados del tratamiento de datos, mecanismos de vigilancia y sanción por mal uso de datos, creación de registro nacional de base de datos y regulación de transferencias a terceros de la información.

El Plan Nacional de desarrollo 2014 – 2018 (Departamento Nacional de Planeación)⁴ le asigno al Departamento Nacional de Planeación la responsabilidad de diseñar la política de explotación de datos (Big Data) para el estado, el cual con esta herramienta puede aportar al cumplimiento de los Objetivos de Desarrollo Sostenible (ODS)

Estos objetivos son:

- Objetivo 9 Industria innovación e infraestructura: (Desarrollo de la innovación en los países con un entorno normativo propicio a las adiciones de valor a los productos básicos
- Objetivo 16 Paz, Justicia e Instituciones sólidas: Promoción del acceso público a la información.

⁴ Bases del Plan Nacional de Desarrollo 2014-2018.

- Objetivo 17 Alianzas para lograr los objetivos: Aumento significativo de datos oportunos, fiables y de alta calidad.

Según el Documento Conpes 3920 a nivel nacional y se espera (Villarrubia, 2018):

- Pasar del 51% de promedio de activos públicos digitalizados y publicados que se tenía en 2017 al 100%.
- Impulsar el porcentaje de entidades que tienen al menos un proyecto de aprovechamiento de datos, que en 2017 era el 9,3% y para 2022 se espera alcanzar el 90%.
- Lograr que por lo menos el 50% de las entidades públicas desarrollen proyectos de aprovechamiento de datos para mejorar los servicios prestados a la ciudadanía.

Los ejes que requeridos para la explotación de estos datos son (Arias, 2018):

1. Generación de datos digitales (insumo),
2. Cultura de datos (demanda),
3. Capital humano para la explotación de datos (oferta)
4. Marco jurídico, ético e institucional de la entidad. Este busca habilitar la generación de valor y reforzar la protección de los individuos en el contexto de disponibilidad y explotación masiva de datos.

Tenemos que tener en cuenta que el 87.3% de las entidades del país se encuentra en un panorama de adaptación si avances a las condiciones mínimas para poder realizar la explotación de los datos que se manejan. Las entidades donde se muestran un

avance significativo más no indispensable son en:

- DNP
- FOGAFIN
- ICFES
- IDEAM
- Súper Sociedades

Actualmente solo se tienen los insumos que se requiere para alimentar la DATA, por lo cual estos serán el punto de partida para crear los datos digitales y tener un punto de partida efectivo. En promedio el 56 % de los activos de información pública son digitales y el 51% de estos están publicados en la web⁵.

“Según la Dirección Nacional de Planeación se mantiene una baja interoperabilidad de las entidades estatales registrando y solo el 4% tiene interoperabilidad. Para entender mejor este concepto la Interoperabilidad se define como la habilidad de dos o más sistemas o componentes para intercambiar información y utilizar la información intercambiada para lograr un fin”

La interacción entre las entidades implica que las organizaciones involucradas compartan información y conocimiento a través de sus procesos de negocio, mediante el intercambio de datos

Un caso claro en que podemos tomar como referencia es el caso del gobierno Perú, y su Proyecto de Gobierno Electrónico, que luego se llamó a partir del año 2009,

⁵ Consejo Nacional de Política Económica y Social, Políticas de explotación de datos Big Data documentos Conpes 3920 – Dirección Nacional de Planeación (2018)

Plataforma de Interoperabilidad del Estado Peruano - PIDE, basado en la Arquitectura SOA, teniendo como proyecto piloto un servicio denominado Constitución de Empresas en Línea, en la que intervienen cinco entidades públicas, actualmente en funcionamiento. El proyecto se ejecutó entre los años 2007 y 2011, es un proyecto modelo de interoperabilidad de éxito a estudiar, por los actores involucrados (instituciones públicas, organismos internacionales, ciudadanos beneficiarios), estrategias empleadas, liderazgo político y técnico, equipos multidisciplinarios, entre otros⁶.

Los inconvenientes que se tiene en la adaptabilidad a crear herramientas necesarias para adoptar de manera adecuada la implementación de la Big Data por medio de la interoperabilidad. Por esto se evidencia en el exceso de trámites y la complejidad de requisitos para acceder a datos públicos.

Las barreras para disponer e intercambiar datos en el sector público son:

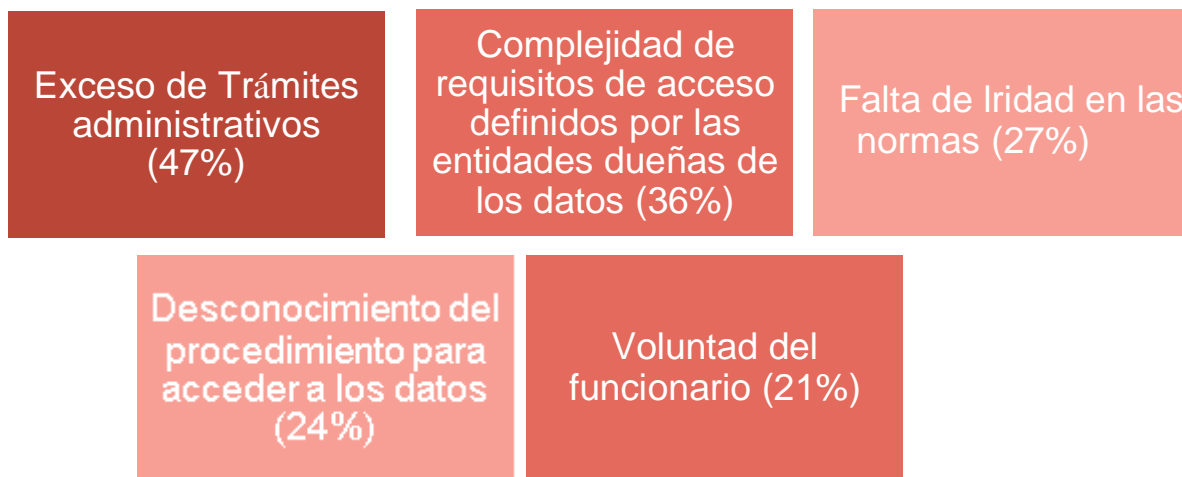


Figura N° 3 Barreras para poder Intercambiar datos en el Sector Publico, Encuesta DNP (2017), Elaboración propia. Universo: 209 entidades (orden nacional). Muestra: 150 entidades

⁶ La Oficina Nacional de Gobierno Electrónico e Informática de la PCM (ONGEI) Premia a Entidades Públicas en Buenas Prácticas Sobre Gestión de Geo- Información <https://www.geoidep.gob.pe/>

CAPITULO 3

COMO LA BIG DATA APOYA LOS PROCESOS INTERNOS DEL MINISTERIO DE JUSTICIA Y DEL DERECHO

La Big Data en el Ministerio de Justicia aprovechara la cantidad de información de tema Jurídico que se exponen a diario tanto de reformas públicas como de las solicitudes que se tienen del ciudadano por lo cual el Ministerio de Justicia y del Derecho tiene como objetivo, dentro del marco de sus competencias, formular, adoptar, dirigir, coordinar y ejecutar la política pública en materia de ordenamiento jurídico, defensa y seguridad jurídica, acceso a la justicia formal y alternativa, lucha contra la criminalidad, mecanismos judiciales transicionales, prevención y control del delito, asuntos carcelarios y penitenciarios, promoción de la cultura de la legalidad, la concordia y el respeto a los derechos; la cual se desarrollará por medio de la institucionalidad que comprende el Sector Administrativo.

La entidad coordinará las relaciones entre la rama Ejecutiva, la Judicial, el Ministerio Público, los organismos de control y demás entidades públicas y privadas, podremos aprovechar la Big Data para el desarrollo y consolidación de la política pública en materia de justicia y de derecho. El Ministerio de Justicia y del Derecho está estructurada de la siguiente forma de acuerdo a la utilización de datos (Ministerio de Justicia y del Derecho, s.f.):

ORGANIGRAMA MINISTERIO DE JUSTICIA Y DEL DERECHO

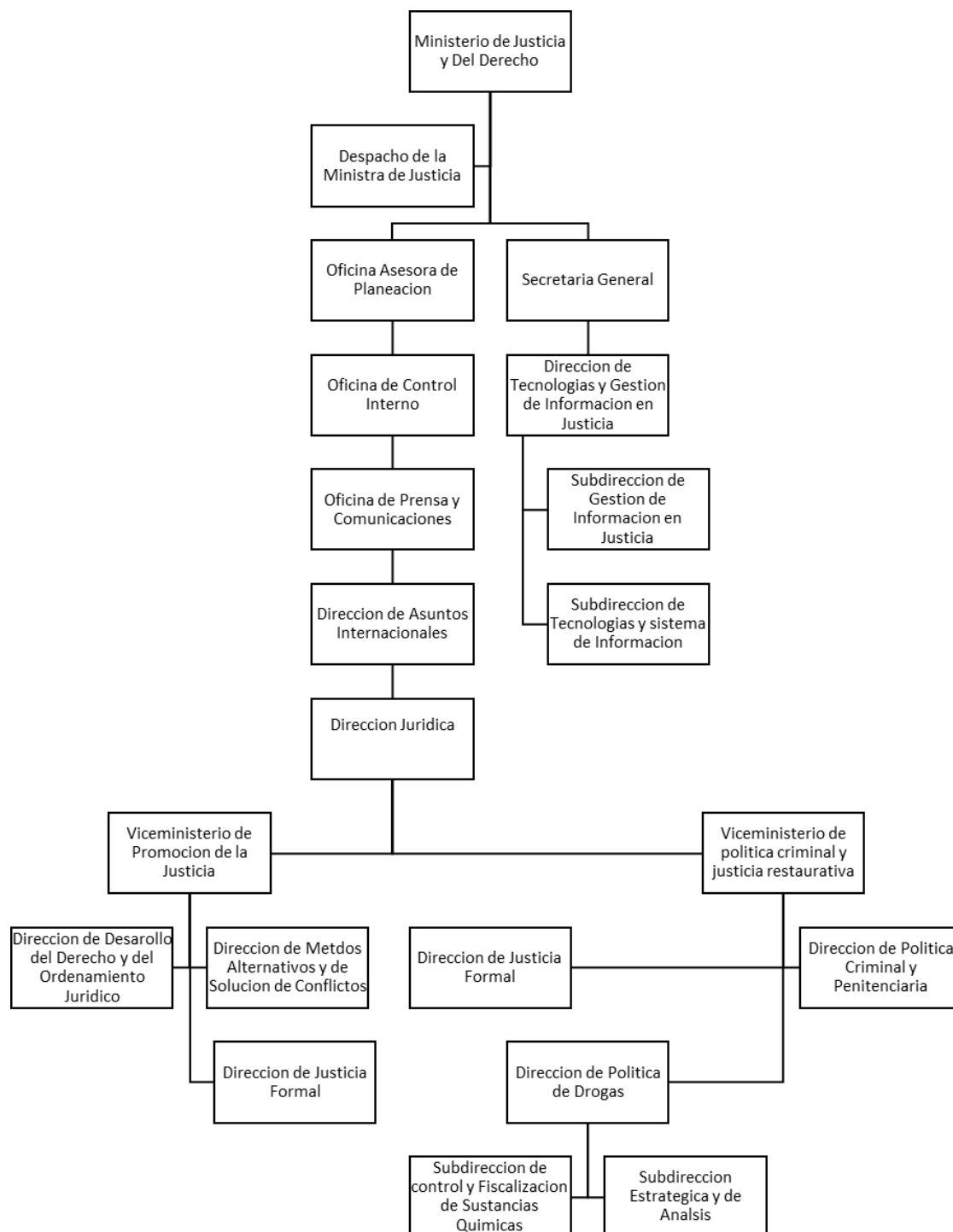


Figura N° 4 Pagina web Ministerio de Justicia y del Derecho/organigrama

DESPACHO MINISTRO DE JUSTICIA Y DEL DERECHO

1. Formular políticas en temas de competencia del Ministerio
2. Coordinar la formulación y seguimiento a las políticas de planes y proyectos del sector administrativo.
3. Presentar proyectos de acto legislativo y de ley ante el congreso de la Republica
4. Implementar, mantener y mejorar el Sistema Integrado de Gestión Institucional.
5. Presidir el consejo nacional de Política Criminal y el Consejo Nacional de Estupefacientes.

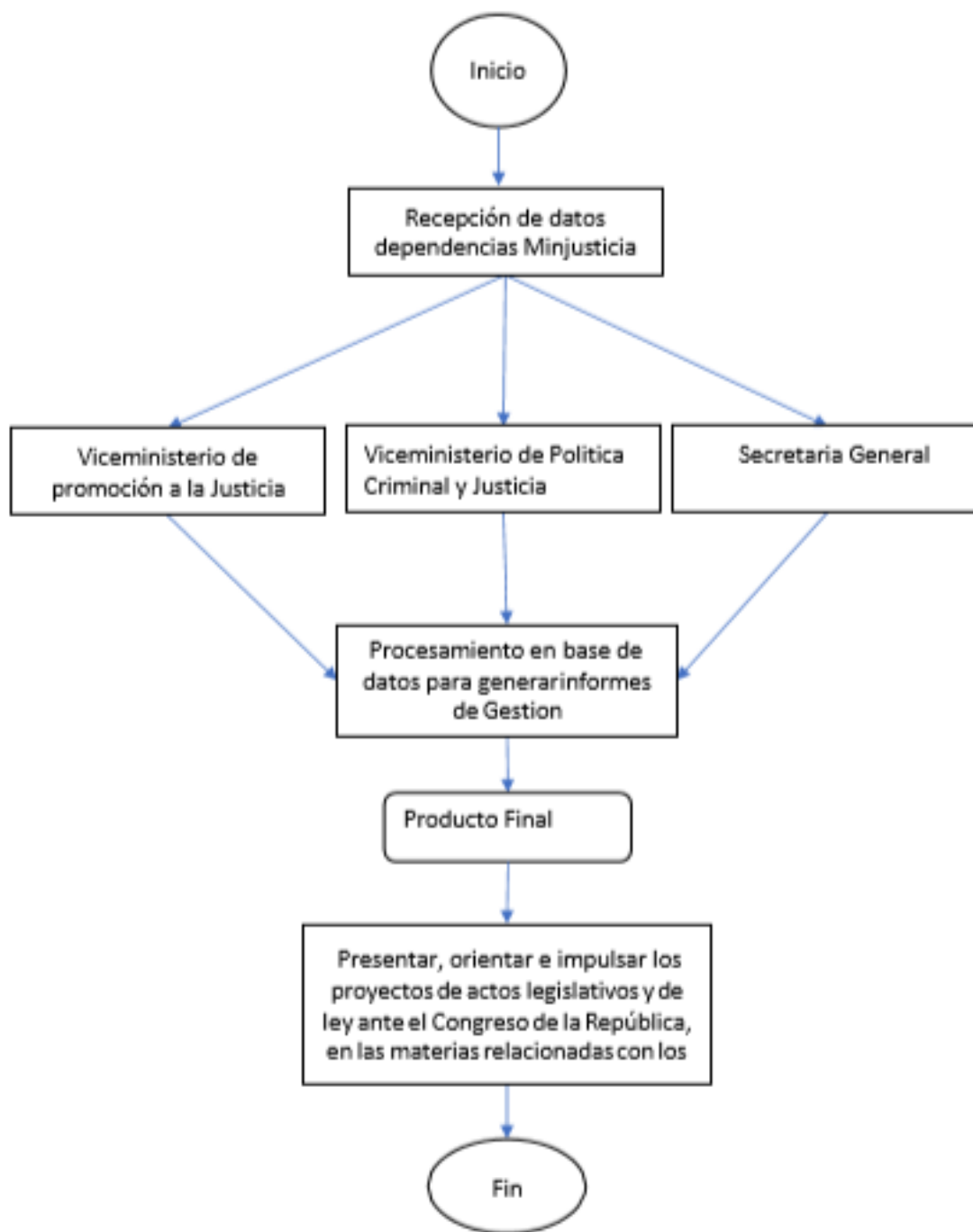
FLUJOGRAMA DE PROCESOS

Figura No. 5 Flujograma de Procesos Big Data, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co

La función principal que tiene el Ministerio de Justicia es formular las políticas públicas en materia de Justicia del derecho, en este caso la utilización de la Big Data en la entidad prácticamente será así:

VICEMINISTERIO DE PROMOCIÓN DE LA JUSTICIA

El objetivo principal del Viceministerio de promoción de la Justicia, es Materializar los derechos de los ciudadanos mediante el fortalecimiento del Estado Social de Derecho⁷. Una de sus funciones más importantes es:

1. Asesorar al ministro en la formulación de políticas y trazar directrices, conjuntamente con el ministro, que orienten el sistema jurídico y su ordenamiento, así como la prestación de los servicios de justicia formal y alternativa, a través de la Rama Judicial, las autoridades administrativas y los particulares.
2. Promover la creación de espacios para acciones en materia de justicia y del derecho, que involucren a la comunidad en la resolución de los conflictos mediante la utilización de los mecanismos formales y alternativos de justicia.
3. Asistir al ministro en el ejercicio de las funciones del Despacho del Ministro de Justicia y del Derecho, en los temas de diseño de políticas, iniciativa legislativa, investigación y análisis de las materias relacionadas con la justicia, propias de su competencia.

⁷ Ministerio de Justicia y el derecho/Organigrama Viceministerio de promoción a la justicia Misión www.minjusticia.gov.co/organigrama.



Figura No. 6 Flujograma de Procesos Big Data Viceministerio Promoción a la Justicia, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co

VICEMINISTERIO DE POLÍTICA CRIMINAL Y JUSTICIA RESTAURATIVA

Con base en el objetivo general del Ministerio de Justicia y del Derecho, son objetivos del Despacho del Viceministro de Política Criminal y Justicia Restaurativa, dentro del marco de sus competencias, formular, adoptar, dirigir, coordinar y ejecutar la política pública en materia de lucha contra la criminalidad, mecanismos judiciales transicionales, prevención y control del delito, asuntos carcelarios y penitenciarios, sus funciones fundamentales son:

1. Asesorar y apoyar al ministro en la formulación y adopción de la política pública criminal, y de manera especial la relativa a asuntos penales y penitenciarios: de drogas, corrupción, trata de personas, crimen organizado y justicia transicional.
2. Asistir al ministro en el diseño de políticas públicas, iniciativa legislativa, investigación y análisis normativo relacionado con la política criminal, las drogas y justicia restaurativa.
3. Planear, coordinar y proponer políticas públicas y trazar directrices conjuntamente con el ministro que orienten la prevención y el control del delito con enfoque diferencial.



Figura No. 7 Flujoograma de Procesos Big Data Viceministerio Política Criminal y Justicia Restaurativa, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co

ESTRUCTURA OFICINAS DE APOYO MINISTERIO DE JUSTICIA

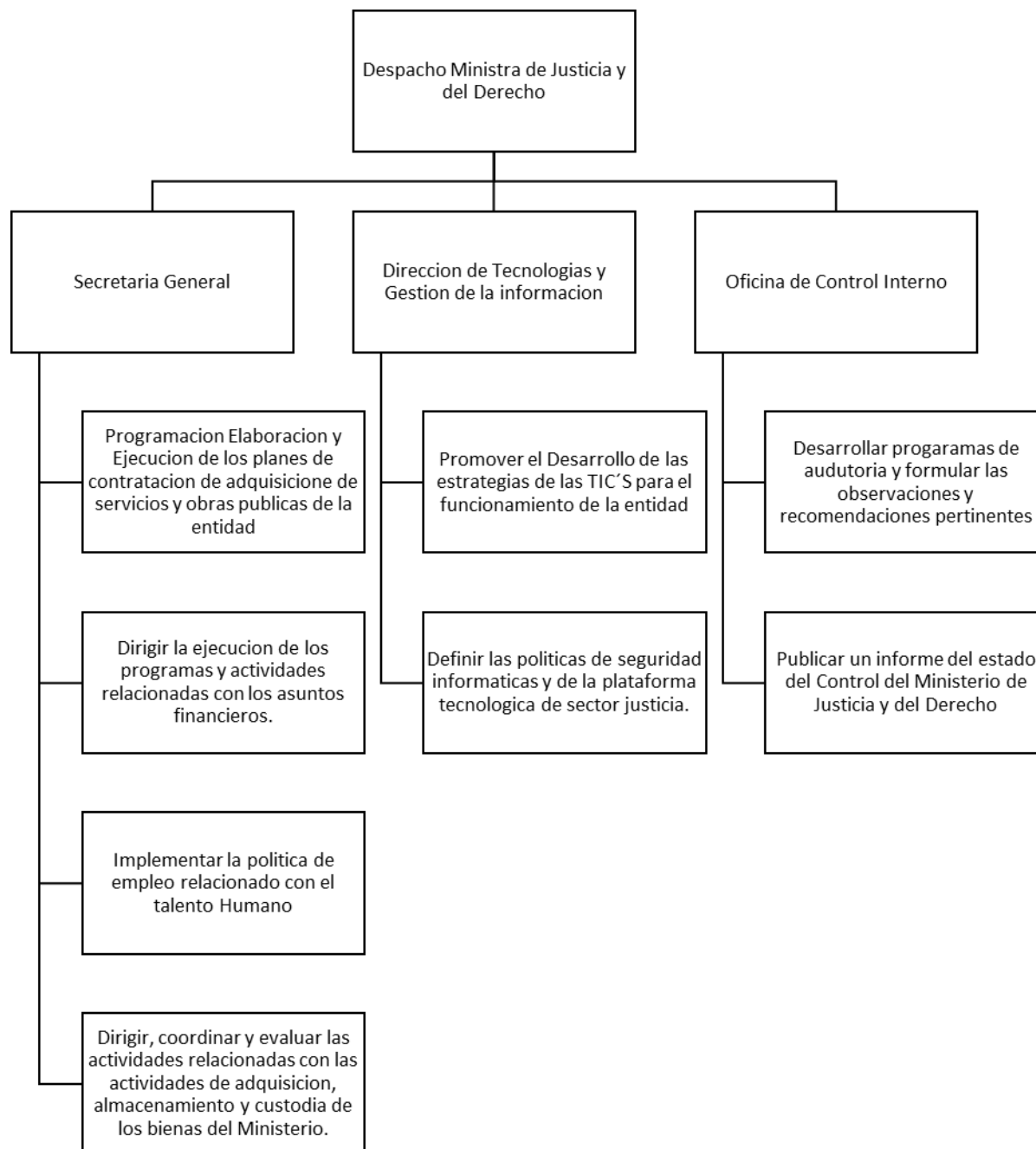


Figura No. 8 Flujoograma de Procesos Big Data Oficinas de apoyo, Ministerio de Justicia y del Derecho Organigrama www.minjusticia.gov.co

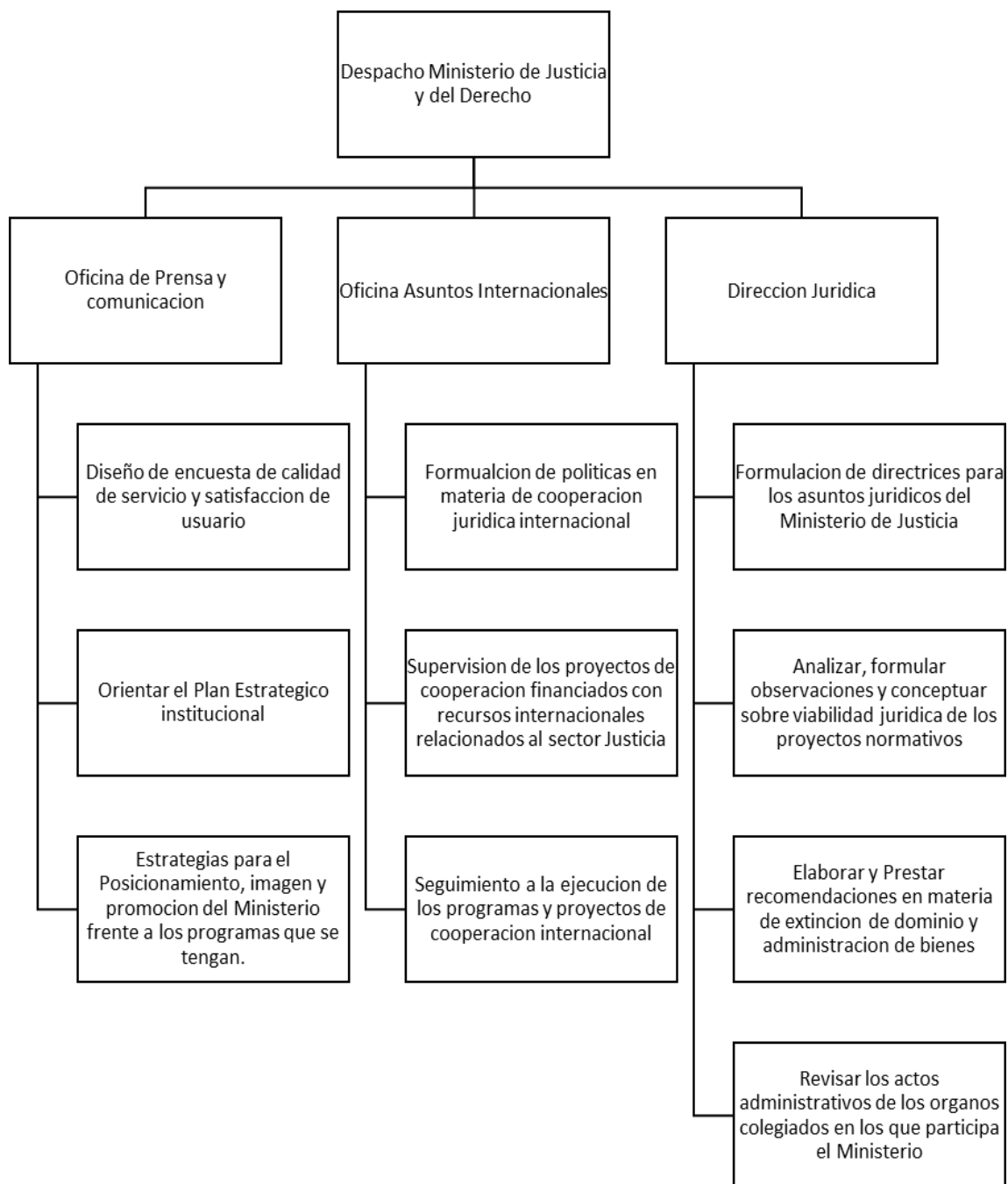


Figura No. 9 Flujograma de Procesos Big Data Oficinas de apoyo, Ministerio de Justicia y del Derecho Organigrama

CAPÍTULO 4

COMPONENTES PARA LA PLATAFORMA BIG DATA DEL MINISTERIO DE JUSTICIA Y DEL DERECHO

En la actualidad disponemos de varios programas para poder implementar la plataforma Big Data con software modernos, de igual manera una plataforma se rige por tres fases principales:

1. **Transformación de Datos:** Primero que todo debemos identificar y unificar los datos que queremos poseer y de los que dispongamos en un único formato y almacenarlos en una base de datos para esto podremos requerir herramientas denominadas ETL que en sus siglas en ingles quiere decir Extract, Transform and Load¹¹, y quiere decir extraer, Transformar y Cargar información (Valero, 2016) que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos y cargarlos en otra base de datos para ser analizados para apoyar la gestión de la entidad. Lo primero que se debe cuando se utiliza procesos ETL es extraer los datos desde el sistema de donde provienen. Esta extracción de datos puede venir en diferentes formatos distintos por lo cual la extracción convertirá los datos a un formato preparado para iniciar el proceso de transformación y no tengan inconvenientes cuando este los reclute. Una desinformación proceso de recolección de datos es la de analizar cada uno de los datos obtenidos, de lo que resulta una verificación si cumplen o no con los datos pautados para la Data de lo contrario serán rechazados.

ESTRUCTURA DE UNA HERRAMIENTA ETL

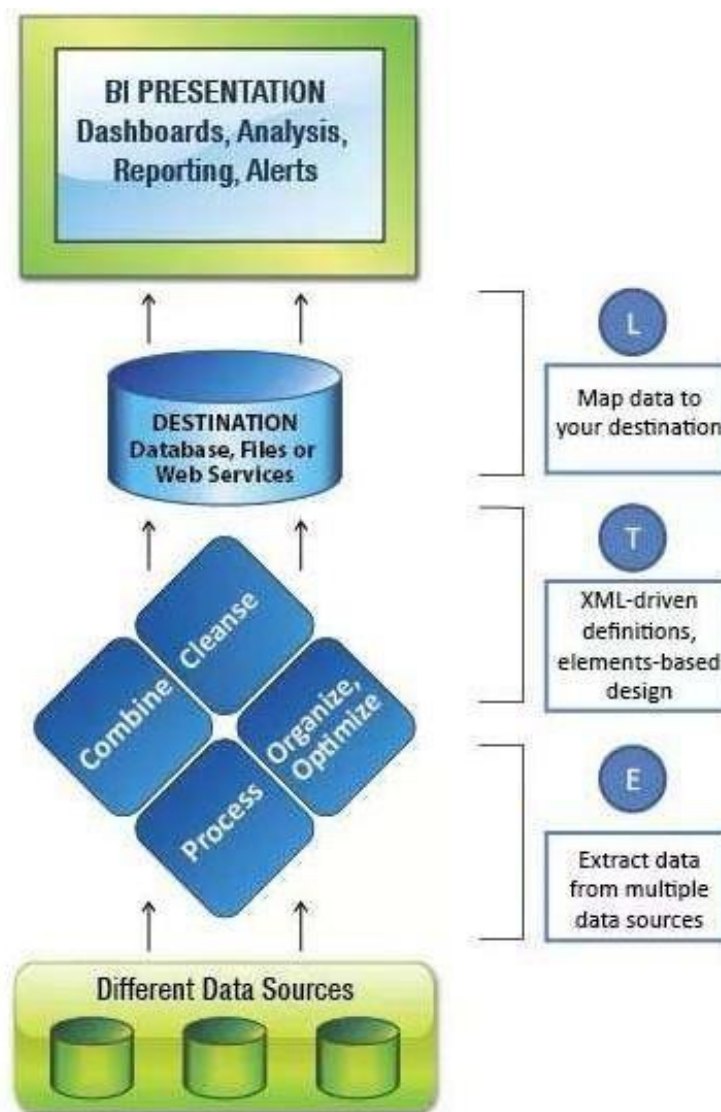


Figura No. 10 Estructura de una Herramienta ETL, Data Prix Herramientas ETL. ¿Qué son, para que valen? Productos más conocidos. ETL's Open Source.

2. ALMACENAMIENTOS DE DATOS:

Para realizar el almacenamiento de este tipo de datos estructurados, no estructurados o semiestructurados, no son posibles las bases de datos

convencionales. Las bases de datos que debemos usar para este tipo de datos son las bases de datos NoSQL⁸ (Not Only SQL) (ACENS) Estas bases de datos almacenan información que no cumplen esquema entidad/relación de las bases de datos relacionales ni utilizan una estructura de almacenamiento en formas de tabla, para almacenar los datos utilizan otro tipo de acciones por las cuales clasificamos estas bases de datos NoSQL. Existen cuatro tipos de bases NoSQL de las cuales solo indicaremos las características principales:

- Almacenamiento clave-valor: Se identifica cada elemento con una clave única lo que permite la recuperación de información de forma muy rápida. Los datos son independientes y no son interpretados por el sistema.
- Almacenamiento documental: Se almacena la información como un documento al que se le asigna una clave única. Es parecida a las bases de datos clave-valor con la diferencia de que además de poder realizar el mismo tipo de búsqueda que en la anterior, en este almacenamiento los datos pueden ser interpretados por el sistema ya que usaremos datos semiestructurados y se realizan consultas más avanzadas sobre el contenido del documento.
- Almacenamiento en grafo: Se basa en la teoría de grafos, la información se representa como nodos y sus relaciones con las aristas de este, para entender esta base de datos pondremos el ejemplo de

¹¹ Acenswhitpapers (2014), Bases de Datos NoSQL que son y tipos que nos podemos encontrar, <https://www.acens.com/wp-content/images/2014/02/bbdd- nosql-wp-acens.pdf>

Facebook, cada nodo sería un usuario el que puede estar relacionado con publicaciones o con otros usuarios por aristas.

- Almacenamiento orientado a objetos o columnas: Almacena columnas en familias y guarda la información de manera cronológica de la misma manera que son representados los lenguajes de programación orientados a objetos como JAVA o Visual Basic.

3. ANÁLISIS DE DATOS.

Ya almacenados los datos recibidos, es necesario diferentes técnicas de análisis por lo cual se tendrán 4 tipos de análisis:

- Por Asociación: Con ello encontramos relaciones entre las diferentes variables utilizando la causalidad. El objetivo es encontrar comportamientos predictivos de otras variables. Para ello, utilizando la asociación se entrega información respecto a la frecuencia con la que se relacionan dos o más atributos de la base de datos.
- Por Minería de datos: Utiliza análisis matemático para deducir patrones y tendencias que existe relacionando los datos entre sí en grandes bases de datos. Integra métodos de estadística aprendizaje automático e inteligencia artificial para lograr los algoritmos predictivos.
- Por Agrupación: Se realiza utilizando el método de clústeres para

segmentar grandes grupos de dato en grupos más pequeños y así poder establecer y descubrir relaciones entre estos subgrupos de datos para poder he intentar predecir el comportamiento de estos.

- Por análisis de texto: Muchos de los datos generados por las personas son textos, ya sea por publicaciones en foros, redes sociales, búsquedas web o e-mails. Este tipo de análisis se utiliza para predecir palabras o extraer información de estos contenidos a gran escala.

Cuando ya se tenga el proceso para poder tomar los datos y analizarlos, el próximo paso será visualizar los resultados en una plataforma que ayude a la entidad en reflejar esta, estos podrían ser los candidatos a establecer el software que dé inicio a la Big Data del Ministerio:

a) **HERRAMIENTA APACHE HADOOP**

Esta plataforma es muy utilizada y ha servido de modelo para el resto de herramientas que realizan la misma labor, Hadoop es un framework que permite procesar grandes volúmenes de datos en lote usando modelos de programación simple. Es compatible con varios servidores, es compatible con diferentes sistemas operativos y se usa frecuentemente sobre cualquier plataforma en la nube como es Amazon, google colud o icloud, además utiliza modelos de programación simples para el almacenamiento y procesamiento distribuidos en datos que se encuentran en clusters.

Apache permite el procesamiento distribuido de grandes conjuntos de datos en varios datos de ordenadores, pero que a ojos del usuario parece un único ordenador. Hadoop separa y distribuye automáticamente los archivos que contienen los datos, además de dividir el trabajo en tareas más pequeñas y ejecutarlas de manera distribuida y recuperarse de posibles fallos automáticamente y de forma transparente al usuario. El proyecto Apache Hadoop consta de los siguientes módulos:

El modelo que se presentara es un modelo de programación que permite distribuir los datos basados en Java pero que también puedan utilizar otros lenguajes de programación avanzados. Básicamente consiste en un mapeo de datos y en barajar los datos más importantes según la programación:

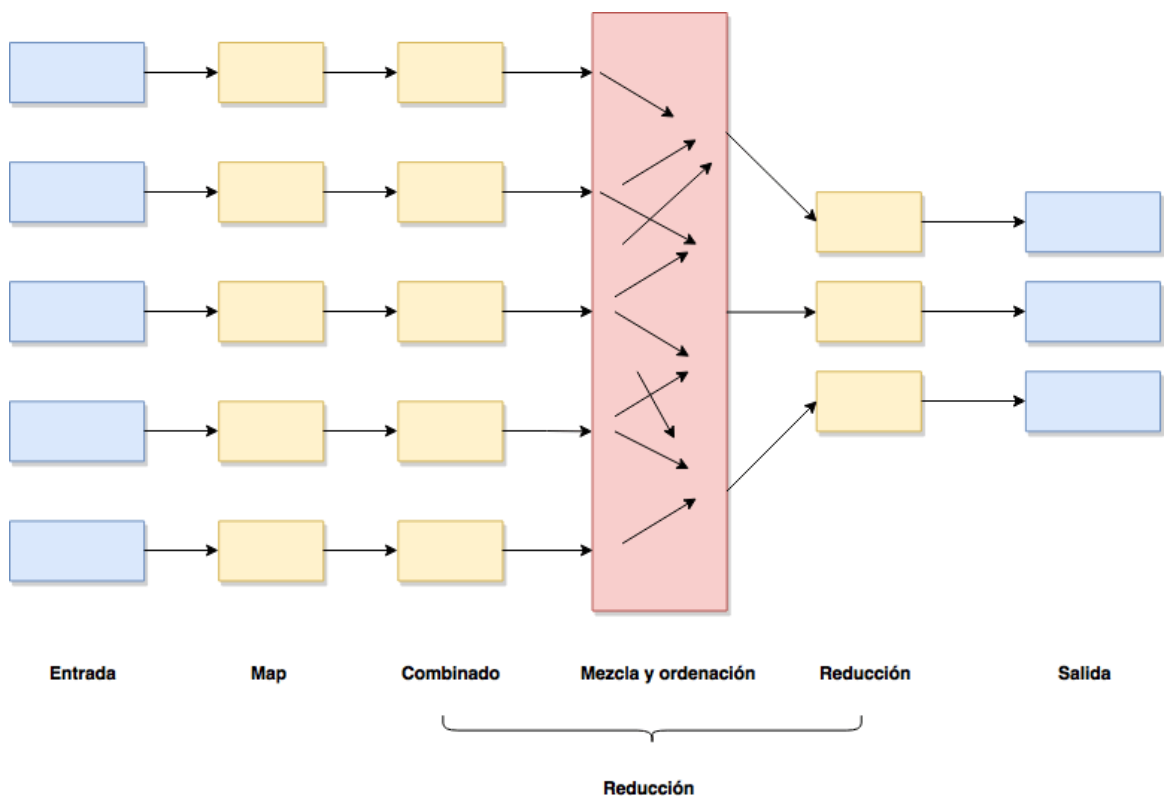


Figura No. 11 Tratamiento de Datos Fundamentos de Apache Hadoop y MapReduce <https://geekytheory.com/fundamentos-de-apache-hadoop-y->

Básicamente lo que se busca es obtener diferentes palabras de un texto y número de veces en las que se repiten lo cual se vinculara más fácil los resultados que se requiere consultar y así ayuda a identificar la información necesaria para los diferentes informes que se requiere en las dependencias.

b) HERRAMIENTA ELASTICSEARCH

Es una potente herramienta para la búsqueda entre grandes cantidades de datos, especialmente cuando los datos son de tipo complejo, permite analizar en tiempo real una gran cantidad de datos y hacer consultas sobre estos mismos. A diferencia de otros sistemas parecidos, no necesita declarar un esquema de la información que añadimos, no sabemos exactamente qué forma van a tener los datos. También podremos hacer visualizaciones del estado de nuestros nodos y escalar sin demasiada necesidad, si se diera el caso.

La herramienta sirve para almacenar, indexar, filtrar y mostrar datos. Está diseñada para continuar su funcionamiento ante la falla de equipos, ya que se puede organizar como una «red de nodos». Es decir, una misma base de datos puede estar replicada en múltiples computadoras conectadas en red, y los datos contenidos pueden estar compartidos en todos los nodos, ya sea de forma íntegra o en pedazos a los que se denomina data shards (OCDS) (Vega , Szyszlican, & Borregón, s.f.).

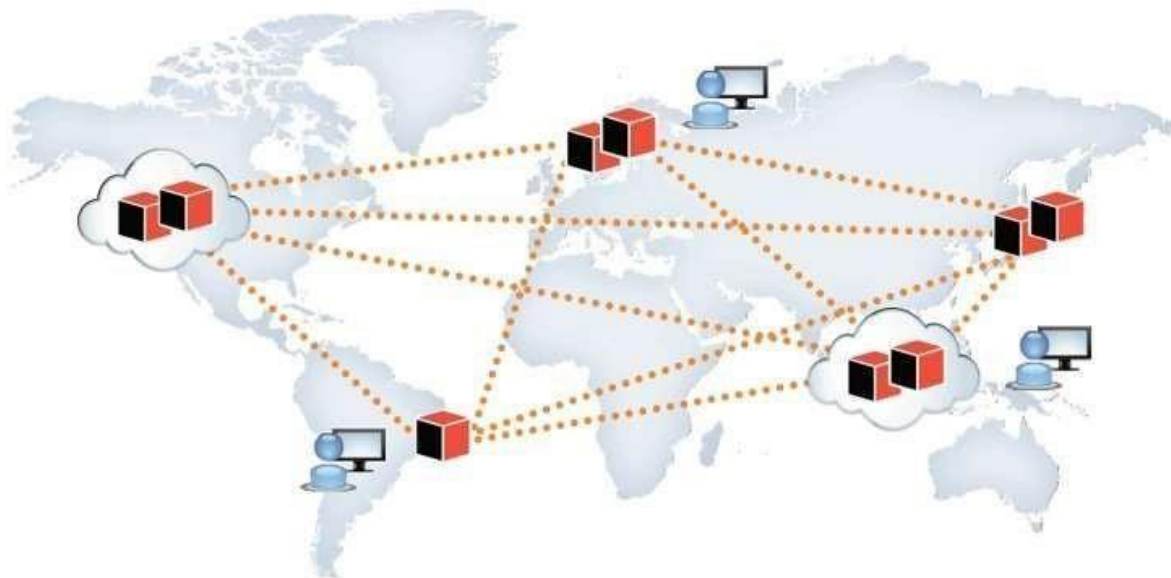


Figura No. 12 Base de datos distribuida Elasticsearch, Importación y análisis de contrataciones en formato Estándar de Datos de Contrataciones Abiertas (OCDS) con herramientas Elastic,

La arquitectura del Elasticsearch permite que la base de datos siga trabajando y captando datos aun cuando uno o varios de los computadores o servidores receptores que componen la red o nodos.

Al conjunto de nodos que se hace en mención, se le conoce como clúster (D'Adamo, 2018)⁹. Un clúster puede tener uno o más nodos los cuales son servidores de Elasticsearch y normalmente estos pertenecen a una sola computadora o servidor. Un clúster tendrá un nombre único, y los nodos reconocerán el clúster al que pertenecen por este nombre y es completamente válido tener un clúster con un solo nodo, esto resulta mucho más sencillo de configurar la recepción de los datos. La herramienta

⁹ Lucia D'adamo, (2018), *Elasticsearch: qué es, cómo funciona y características*, <https://www.consulthink.it/es/elasticsearch-que-es-como-funciona-y-caracteristicas/>

distribuye la información y también distribuye el procesamiento de consultas de datos que se requieran. Cuando se realiza una búsqueda en la base de datos y esa información se encuentra distribuida en varios ordenadores, será tarea de cada nodo procesar dicha consulta y devolver la información que tenga disponible e indispensable. Al final, el sistema en su conjunto creará un resultado final conjuntando los segmentos devueltos por cada nodo.

Esto permite que un sistema basado en Elasticsearch y configurado apropiadamente pueda realizar consultas complejas, en grandes cantidades de datos conservando un excelente rendimiento para el usuario final, más allá de su ubicación.



Figura No. 13 Elasticsearch Cluster, Importación y análisis de contrataciones en formato Estándar de Datos de Contrataciones Abiertas (OCDS) con herramientas Elastic,

Los componentes de índice para los datos encontrados se denominan INDEX y funciona principalmente como un viejo índice de biblioteca donde se encuentran las referencias de las búsquedas realizadas donde comparten características tanto de nombres de los datos como, direcciones, teléfonos, números de procesos administrativos etc. A estos índices se pueden hacer consultas, actualizaciones, borrado o agregado de datos usando el nombre correspondiente.

Un índice puede contener mucha más información de la que es posible almacenar físicamente en una sola computadora o de la que sería eficiente procesar en una sola computadora. Para solucionar este problema, se puede subdividir los índices en fragmentos más pequeños llamados «shards». Esta función es la que le permite ser eficiente con cantidades de datos monumentales y puede subdividir el índice permitiendo que se puedan agregar más nodos al clúster y los documentos pueden estar distribuidos entre todos los nodos.

CAPITULO 5

ESTRATEGIA PARA APROVECHAR LA BIG DATA EN EL MINISTERIO DE JUSTICIA Y EL DERECHO

La Big Data en el Ministerio de Justicia ayudara a tratar en un manejo adecuado a recolectar una concentración grande de volumen de datos preciados llenos de información para la actividad de la entidad lo cual nos ayudara a tomar decisiones analizando cada uno de los datos recolectados. En la actualidad el Ministerio solo se tienen los insumos que se requiere para alimentar la DATA (datos estadísticos), por lo cual estos serán el punto de partida para crear los datos digitales y comenzar a trabajar. Actualmente en promedio el 56 % de los activos de información pública son digitales y el 51% de estos están publicados en la web, esta información esta para consulta de las personas que requiera¹⁰n acceder a esta información (Arias, 2010).

Primero que todo debemos definir los datos que genera el Ministerio de Justicia y su volumen. Por lo cual reflejaremos el punto de inicio de la información que podemos aprovechar para crear una Big Data y estos son:

¹⁰ Diana Arias, (2018), Colombia es el primer país en explorar la Big Data, <https://www.enter.co/especiales/empresas/colombia-primero-explotacion-de-big-data/>

Tabla 2 Mapa Estratégico Cultura de Datos

ESTRATEGIA	FUNCIONALIDAD	PROCESO
<p>Recopilación de Datos de Personas Naturales o Jurídicas en consultas externas</p>	<p>En esencia la recopilación de datos personales nos será útil para crear un intervalo o filtro de las peticiones que se tiene en casos de consultas o peticiones por temas legales al Ministerio el cual servirá para identificar de manera más ágil y rápida el tema relacionado con el tercero. La acción que utilizaremos será el registro de la información de los usuarios en las bases de datos del Ministerio de Justicia y del Derecho, con la finalidad de analizar,</p>	<p>Estos datos serán recolectados por medio de los canales presenciales de recepción de documentos o solicitudes que serán recibidos en las instalaciones físicas de la entidad y también por medio de correo electrónico el cual estará disponible para atender cualquier solicitud y así aprovechar los datos obtenidos</p>

	<p>evaluar y generar datos estadísticos, así como indicadores sectoriales para la formulación de políticas en el sector</p>	
<p>Consultas Plataforma Legal App</p>	<p>Permite a los usuarios que necesiten conocer cómo adelantar un trámite o hacer uso de algún servicio relacionado con la Justicia. En términos generales se digita una palabra clave según el caso requerido y esta herramienta orienta a las personas sobre qué hacer, la autoridad o institución a la cual puede acudir y la ubicación exacta en su municipio.</p> <p>Esto nos permite orientar más a la Data y comprender más las</p>	<p>Se medirá las consultas diarias a la plataforma y se recopilará en la Data las búsquedas más frecuentes para así fortalecer los temas más frecuentes y darle un valor agregado en temas legislativos y jurídicos para el ciudadano.</p>

	<p>necesidades de búsqueda de los usuarios y permitir al Ministerio de Justicia mejorar en las políticas que se han consultado y así brindar un mejor servicio judicial al ciudadano.</p>	
<p>Recopilación de los procesos internos de las dependencias</p>	<p>Genera un valor agregado para la Entidad ya que permite tanto al ministro como a las demás dependencias conocer y empalmar diferentes trabajos o productos realizados por las dependencias y así crear informes más completos y siendo más eficiente en los tiempos de ejecución de estos ya que facilita la intercomunicación entre oficinas.</p>	<p>Se genera el producto final de cada dependencia el cual será cargado en un servidor o unidad y esta se relacionará en la DATA con un nombre específico para su fácil búsqueda para las demás dependencias o para el Ministro de Justicia, permitiendo claridad y agilidad en la información.</p>

En el caso de la Recopilación de Datos de Usuarios, los datos personales donde recaen la información primaria de contacto lo cual permitirá recolectar, almacenar, procesar, usar y transmitir o transferir (según corresponda), lo cual permita identificar la línea de tendencia donde venga la solicitud y permita una mejor brevedad en la respuesta según requerimiento previo.

Estos datos recolectados al ser de terceros o usuarios se manejarán de forma estricta según los deberes de seguridad y confidencialidad ordenados por la Ley 1581 de 2012 y el Decreto 1377 de 2013, compilado en el Decreto 1074 de 2015 (Decreto Único del Sector Comercio, Industria y Turismo).

CAPITULO 6

SEGURIDAD EN LA EXPLOTACIÓN DE LOS DATOS.

La importancia de la información que se maneja en el Ministerio será, en gran medida, relativa a nuestro sector justicia, así:

- El Ministerio de Justicia maneja un gran volumen de información personal de personas de tipo naturales como jurídicas, a la que se deben aplicar todas las medidas de seguridad para evitar que se pierda, modifique o se acceda a ella sin autorización. Además, suele ser necesario llevar un registro de los accesos y modificaciones si llegara el caso.
- Algunas peticiones, quejas, reclamos o simplemente documentación que se obtienen son de carácter confidencial tanto cuya difusión puede suponer una importante pérdida que puede perjudicar a la nación en temas jurídicos.

El riesgo en la recolección y aprovechamiento de datos no se concentra solamente en el uso del dato, sino también en la conversión del mismo en información valiosa para la organización, tanto para la estructuración de procesos de negocio como para el acercamiento efectivo y eficiente a los clientes. Por eso mismo el Ministerio de Justicia tiene una política de tratamiento sobre los datos ordenados por la Ley 1581 de 2012 (DiazGranados Ordóñez, 2016) y el Decreto 1377 de 2013, compilado en el Decreto 1074 de 2015 (Decreto Único del Sector Comercio, Industria y Turismo).

En primer lugar, se debe revisar qué información tratamos (bases de datos, archivos, aplicaciones, programas,) y seleccionaremos la más crítica, la que, si nos faltara, por su confidencialidad o si se corrompiera, paralizaría nuestra actividad y nos acarrearía pérdidas de imagen ante el país. En esta clasificación de la información podemos establecer varios niveles en función de su importancia para el Ministerio de Justicia:

a. CONFIDENCIAL:

Información sensible para el Ministerio. Su acceso está restringido únicamente a la Dirección y a aquellos empleados que necesiten conocerla para desempeñar sus funciones. También datos de carácter personal, en particular los de categorías especiales. Esta información debe marcarse adecuadamente. Se deben implementar todos los controles necesarios para limitar el acceso a la misma únicamente a aquellos empleados que necesiten conocerla. En caso de sacarla de las instalaciones de la empresa en formato digital, debe cifrarse. Para los datos de carácter personal, se deben tener en cuenta la protección y garantías indicadas en la legislación sobre la materia, como se expuso anteriormente.

b. INTERNA

Esta Información propia de la empresa puede ser accesible para todos los empleados. Por ejemplo, la política de seguridad de la compañía, el directorio de personal u otra información accesible en la intranet del Ministerio de Justicia. Esta información debe estar adecuadamente etiquetada, y estar accesible para todo el personal. No debe difundirse a terceros salvo autorización expresa de la Secretaria General.

c. PÚBLICA

Cualquier material de la empresa sin restricciones de difusión. Por ejemplo, información publicada en la página web. Esta información no está sujeta a ningún tipo de tratamiento especial. Un aspecto importante a considerar en la selección e implantación de controles es su tipología o naturaleza. Ésta puede ser:

- Técnica: medidas de carácter tecnológico dentro del ámbito de la seguridad, puede ser medidas como antivirus, cortafuegos o sistemas de copias de seguridad.
- Organizativa: medidas que se centran en la mejora de la seguridad teniendo en cuenta a las personas, por ejemplo: formación en seguridad, identificación de responsables o implantación de procedimientos formales de alta y baja de usuarios.
- Física: medidas físicas para proteger nuestra organización. Como, por ejemplo, acondicionar adecuadamente la sala de servidores frente a riesgos de incendio, inundaciones o accesos no autorizados, establecer un sistema de control de acceso para entrar en las oficinas, poner cerraduras en los despachos y armarios o guardar las copias de seguridad en una caja ignífuga.

Por defecto, toda organización debe seguir el principio del mínimo privilegio para acceso a la información. Este principio se traduce en que un usuario sólo debe tener acceso a aquella información estrictamente necesaria para desempeñar sus funciones diarias y así generar un protocolo que sirva para aislar e identificar el

personal que mantiene esta información en custodia. Para conseguir este objetivo, previo a la implementación de medidas técnicas, debemos realizar los siguientes pasos:

- Definir los diferentes tipos de información que existen en nuestra organización: datos de recursos humanos, financieros, políticas de drogas, políticas judiciales, sustancias químicas o psicoactivas, información de proceso judiciales, información personal de los ministros, viceministros, documentos contractuales, proveedores, etc.
- Establecer quién puede acceder a cada tipo de información. Para acometer esta tarea puede ser útil, si la estructura organizativa lo permite, realizar una matriz que cruce información con áreas o departamentos que tienen necesidad de acceso a dicha información

De acuerdo al Ministerio de Tecnologías de la Información y las Comunicaciones publica el Modelo de Seguridad y Privacidad de la Información (MSPI) (Ministerio de Tecnologías de la Información y las Comunicaciones, s.f.) el cual estará acorde con las buenas prácticas de seguridad y será actualizado periódicamente; reuniendo los cambios técnicos de la norma 27001 del 2013, legislación de la Ley de Protección de Datos Personales, Transparencia y Acceso a la Información Pública, entre otras, las cuales se deben tener en cuenta para la gestión de la información. Mediante la adopción del Modelo de Seguridad y Privacidad por parte de las Entidades del Estado se busca contribuir al incremento de la transparencia en la Gestión Pública, promoviendo el uso de las mejores prácticas de Seguridad de la Información como base de la aplicación del concepto de Seguridad Digital

CAPITULO 7

CULTURA DE DATOS

Implementar una cultura basada en datos es poner en práctica una serie de cambios en el uso de los recursos y en la mentalidad de los empleados de una empresa para que puedan tomar mejores decisiones con los datos. El modelo deja de lado la intuición la tecnología, como Big Data, Inteligencia Empresarial y Business Analytics, para que funcionen y aumenten las tasas de éxito alcanzadas para el Ministerio de Justicia.

Los beneficios que traerá será el cambio del instinto a los datos lo cual es ventajoso en cualquier área de operación del Ministerio de Justicia y del Derecho. Hoy en día, todas las principales organizaciones del mundo apuestan por Big Data para comprender mejor al consumidor y predecir sus comportamientos y deseos, en el caso del Ministerio será más fácil comprender cuales son las falencias y los deseos de los ciudadanos en temas judiciales que les ayuden en sus vidas y también conocer en temas nacionales que política pública puede ayudar en temas jurídicos.

Algunos beneficios son:

- Actualizaciones de tecnologías utilizadas en el negocio;
- Integración de la gestión con el software en la nube, lo que aumenta la accesibilidad de los datos;
- Cambio en la mentalidad de los empleados y optimización de los procesos llevados a cabo en la rutina diaria.

- Optimización del intercambio de información y toma de decisiones;
- Centrarse en el ciudadano y sus expectativas;
- Identificación de amenazas que comprometan la continuidad de las operaciones judiciales;
- Agilidad para responder a los cambios en la política pública
- Ahora la pregunta importante es cómo podemos implementar la Big Data en los trabajadores:



Figura No. 14 Flujograma de Procesos Organizacional Big Data, MyABCM ¿Cómo puede su empresa ser innovada por la cultura basada en datos?

Paso 1: Conocer cómo obtener y analizar los datos. (MyABCM, s.f.) ¹¹

Es importante para los funcionarios y contratistas del Ministerio de Justicia y del Derecho el método de recolección de datos necesario para identificar los diversos requerimientos que se tengan en sus labores diarias y así como las peticiones que se tengan tanto del gobierno como de las personas naturales.

¹¹ Myabcm, ¿Cómo puede su empresa ser innovada por la cultura basada en datos? <https://www.myabcm.com/es/blog-post/cultura-basada-en-datos/>

De modo que se tiene que tener en cuenta que para recolectar la información vital se requiere lo siguiente:

1. Tener un instrumento confiable para medir los datos recogidos los cuales sean confiables y que sean acertado a los requerimientos solicitados. (Acá entraría las herramientas de Big Data que se mencionaron en el capítulo 4).
2. Ponerlo en práctica.
3. Organizar la información para poder crear informes finales.

Ahora bien, para analizar los datos se deben examinar detalladamente la información arrojada para poder sacar las conclusiones y tomar decisiones para los informes finales de gestión que se tengan en la entidad.

En este caso antes de obtener los datos, se determinará qué información se requiere para construir la información y una vez tengamos los datos correctos entrará el momento de realizar el análisis pertinente de estos mismos, buscando la información que arroje al funcionario de acuerdo al requerimiento solicitado y entregar las bases para proceder con la construcción del insumo final.

Paso 2: Transformar las ideas en decisiones procesables

Este paso permite a la entidad y a los funcionarios tomar la información obtenida y analizada para construir un enfoque más claro en las decisiones que se harán con este insumo y así genere un impacto positivo para la política pública del país en

materia de justicia, pero fortaleciendo principalmente en la defensa del ordenamiento jurídico, gestión jurídica pública del derecho, socialización de la información jurídica que disponga la entidad y un acceso a la justicia para todas las personas que lo requieran

Paso 3: Aliene a sus empleados a que hagan lo mismo

Para implementar este paso se deberá generar una cultura de datos dentro de la entidad lo cual aproveche al máximo los datos que se quieran recolectar. Se deberán crear capacitaciones en los sistemas de gestión que se acoplen a los ya establecidos en los manuales de funciones de los funcionarios y que esto permita un manejo amigable con el software que se valla a utilizar y así usar esta herramienta correctamente para que los empleados siempre obtengan la información factible y necesaria para la entidad.

Paso 4: Capacitación de equipos de trabajo en conocimiento de datos

Para lograr este nivel de captación de datos en el Ministerio se deberá implementar unas estrategias basadas en manejo de datos en el sector público, por lo cual se deberá:

- Definir una Metodología para medir la preparación e implementación en un plan de preparación para la explotación de datos.
- Para tener los insumos necesarios para generar una buena implementación de recolección de la data, se deberá crear un acuerdo marco de precios que

facilite la adquisición de bienes y servicios para lograr este objetivo.

- Por último, se debe incorporar la explotación de datos en la formulación de inversión de la entidad lo cual quedara plasmado en el Plan anual de adquisiciones.

Con todo lo expuesto anteriormente y sin importar el tipo actividad que se realice, hay tres obstáculos comunes al adoptar soluciones analíticas y de Big Data en el Ministerio de Justicia:

- La falta de conocimiento sobre estrategias y métodos para realizar análisis que realmente ayuden a mejorar en la entidad.
- En segundo lugar, la carencia de gestión administrativa para implementar proyectos de analítica debido a la condición de competencia con otros proyectos concebidos como de mayor prioridad;
- La carencia de habilidades en los equipos internos de la organización para efectuar proyectos analíticos complejos.

Al implementar un proyecto de analítica no debe de concebirse como un proceso de difícil desarrollo; todo se basa en la relación adecuada entre los objetivos de la entidad, la estrategia y la tecnología existente, las cuales deben de estar enfocadas a resolver requerimientos y obtener la información más precisa y fidedigna, sin importar el área de la empresa que lo requiera

Por eso mismo uno de los puntos que se debe reforzar para aprovechar al máximo la data de la entidad es el capital humano, por esto se deberá:

- Realizar una medición de la brecha con respecto a la data que se quiere realizar.
- Implementación de una política de entrenamientos en la explotación de datos frente a capacitaciones a la entidad.

Paso 5. Definir la capacidad de acción de las dependencias

Para implementar estas capacidades se deberá:

- Elaborar una guía que garantice la protección de datos personales en las entidades públicas y actualizar unos lineamientos de contratación para incorporar la explotación de datos.
- Articular el modelo de gestión, para la explotación de datos con el modelo integrado de planeación y gestión.
- Formulación de una propuesta de norma para ajustar el marco jurídico de la institución en frente a la implementación de la Big Data.

CAPITULO 8

BALANCED SCORECARD BIG DATA MINISTERIO DE JUSTICIA Y DEL DERECHO

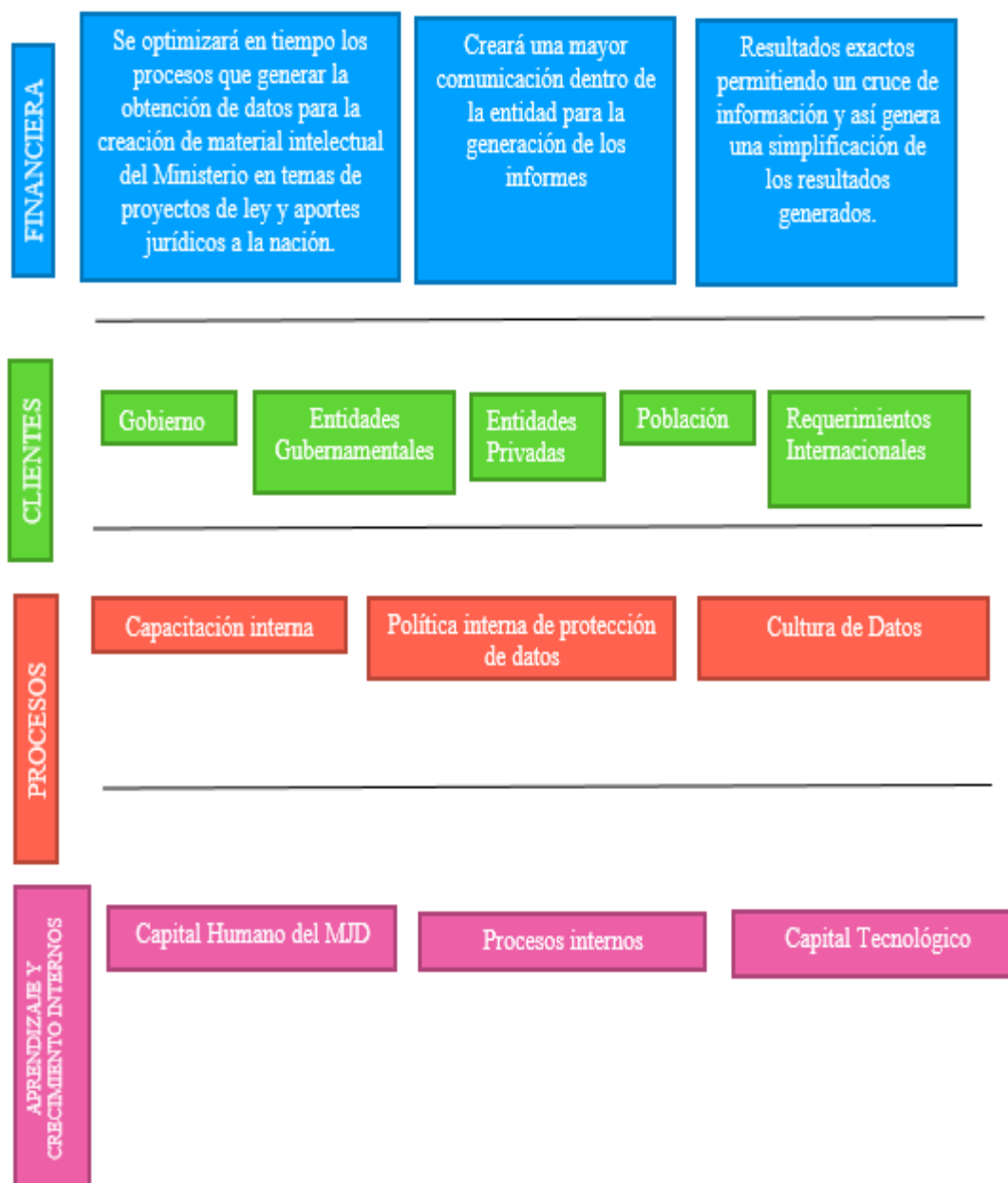


Figura No. 15 Balance ScoreCard Big Data MJD

De acuerdo al cuadro anterior el Balanced Scorecard permite identificar los cuatro aspectos fundamentales, en la entidad para entender en si el conjunto de acciones que puede generar la implementación de la Big Data en el Ministerio de Justicia y del Derecho. Por este motivo se identificará cada aspecto que conforma esta figura:

1. **Aprendizaje y crecimiento interno:** En este aspecto encontramos unido el capital humano, los procesos internos y el capital tecnológico. Todo esto representa el inicio del proyecto Big Data del Ministerio de Justicia y del Derecho. De este punto se toma el personal esencial para obtener los datos requeridos, donde serán orientados con un plan de trabajo que los guíe a cumplir con los requisitos necesarios para obtener la valiosa información dando un valor agregado a la gestión del Ministerio frente a los diversos temas que tenga que ver con el sector justicia de la Nación. Para lograr esto es esencial la ayuda de la parte tecnología, donde entra las herramientas tecnológicas que buscan reducir esa carga exhaustiva de datos, permitiendo que sea más fácil y rápida de acceder la información requerida.
2. **Procesos:** Esta parte se le conoce como la fase de aprendizaje del personal del Ministerio de Justicia, y se caracteriza en la profundización e inducción de los temas operativos correspondientes al software o herramienta informática que permita la recolección de Datos. Esto conlleva a crear una política interna en toda la entidad con respecto a la

protección de datos, la cual permite que la información recolectada será manejada de la mejor forma posible con límites necesarios para su acceso de acuerdo a la información adquirida. Todo esto ayuda a crear una cultura de datos que será esencial para todo el personal que manipule la Data conociendo los procesos que se deben hacer al tener acceso a esta información.

3. **Clientes:** Esta parte de la estructura representa los autores que obtienen los informes finales que se adquieren al recaudar la información con ayuda de la Big Data y que ayudan a dar respuesta a las peticiones que realicen las entidades del Gobierno, Población en General, Entidades Internacionales y demás actores que requieran un concepto Jurídico.

4. **Financiera:** Esta etapa demuestra el mejoramiento y ganancia que tendrá las áreas del Ministerio de Justicia y del Derecho frente a la optimización de los tiempos en dar respuestas a los requerimientos de manera exacta y verídica, dando así un valor agregado a la entidad donde se refleje una comunicación más precisa y fortaleciendo un apoyo continuo hacia al Gobierno Nacional frente a los temas que requieran atender a nivel nacional en temas jurídicos. Esto es considerado como un aspecto positivo puesto que gracias a ese mejoramiento puede que genere un impacto positivo y le den un mayor presupuesto a esta cartera Ministerial para realizar sus diversos proyectos y pueda ayudar a mejorar el País y a sus ciudadanos.

CONCLUSIONES

La Big data ayudara al Ministerio a crear un mecanismo donde pueda agilizar sus procesos internos, guardando información física en respaldos digitales para así poder crear una red de información generosa y que cualquier entidad estatal necesite y pueda solicitarla de manera más rápida y segura enfocada en proceso judiciales o temas que tengan que ver con el sector justicia. La Big Data ayudara a la entidad a dar respuestas más rápidas agilizando los canales de recepción y envío de información.

Esta Data digital ayudara a tener un respaldo más eficiente que el fisco ya que no es necesario desplazarse donde está este material el cual está guardado en archivos físicos, si no que puedan ser consultados por personal autorizado y dar al solicitante. En la red se busca que la información que llegue ayude a la entidad a generar un pronóstico de la situación actual del país la cual genere una política Jurídica a los casos más buscados o consultados por los ciudadanos o por las empresas, este es el fin de la Big Data en el Ministerio de Justicia y del Derecho.

La Big Data por sí misma no sirve de mucho si no hay una estructura que la pueda analizar y organizar, es decir, transformar la información en soluciones a problemas específico. Cada búsqueda que realizamos en la internet, cada publicación en redes sociales y básicamente, cada movimiento que hacemos en el mundo digital, genera información que nos ayudara a comprender más lo que nosotros buscamos o necesitamos. Por su volumen, a esta información le llamamos Big Data y desde hace tiempo su análisis sirve para crear nuevos productos o servicios diseñados a la

medida de las necesidades específicas de los usuarios.

Las recopilaciones de grandes cantidades de datos permiten que las empresas o entidades se muevan mucho más rápidamente, sin problemas y de manera eficiente. También les permite eliminar las áreas problemáticas antes de que los problemas acaben con su credibilidad o su eficiencia. El análisis de Big Data ayuda a las organizaciones a aprovechar y utilizar los datos para identificar nuevas oportunidades. Eso, a su vez, conduce a movimientos de estratégicos más inteligentes, operaciones más eficientes, mayores políticas y usuarios más felices. En esta era digital y de avanzada tecnología, se genera gran cantidad de datos a nivel global, por lo cual el volumen generado diariamente es tan grande que las bases de datos actuales se están quedando obsoletas por lo avanzado que está teniendo la sociedad, por tal motivo los datos crecen más rápido que la capacidad de datos actuales que se almacena la información en la actualidad. Si las entidades estatales no se aprenden a tratar y administrar este tipo de flujo de información dejarán de ser competitivos frente a las necesidades que se tengan los usuarios. Con los tratamientos de datos el Ministerio de Justicia el objetivo principal es tener una red de servicios controlada para brindar correctamente la información solicitada a usuarios, información sobre procesos judiciales contra personas naturales, jurídicas y entidades que tengan índole de urgencia para prevalecer la protección y debido proceso de la nación.

Una de las debilidades de las entidades públicas es que no se tiene una trata de datos adecuado de sus procesos que desempeñan lo cual es un arma de doble filo ya que no somos conscientes de la cantidad de datos que disponemos y normalmente estos datos

están desordenados y desaprovechados

Hoy en día en la era tecnológica, se tienen muchas opciones que nos pueden ayudar a recopilar la información recibida por los usuarios, lo cual generaran la Data necesaria para alimentar nuestro banco de datos y así poder analizar, lo cual nos ayudara a comprender más las necesidades de las personas en temas de Justicia. La Big Data, viene dada porque permite explotar el valor de los datos a los que hasta ahora no se les había prestado atención ni podido extraer ningún valor.

Bibliografía

- ACENS. (s.f.). *Bases de datos NoSQL, Qué son y tipos que nos podemos encontrar*. Obtenido de Acens:
<https://www.acens.com/wp-content/images/2014/02/bbdd-nosql-wp-acens.pdf>
- Arias, D. (2018). *Colombia es el primer país en explotar Big Data en Latinoamérica*. Obtenido de Enter.co: <https://www.enter.co/especiales/empresas/colombia-primer-exploracion-de-big-data/>
- Barranco Fragoso, R. (2012). *¿Que es Big Data?* Obtenido de IBM:
<https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html>
- Blanco, T. P., & Sánchez Herrera, J. (2014). *Nuevas tendencias en comunicación estratégica*. Madrid: ESIC EDITORIAL.
- D'Adamo, L. (2018). *Elasticsearch: qué es, cómo funciona y características*. Obtenido de Consulthink:
<https://www.consulthink.it/es/elasticsearch-que-es-como-funciona-y-caracteristicas/>
- Departamento Nacional de Planeacion. (2018). *Politica Nacional de Explotacion de Datos (Big Data)*. Obtenido de <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3920.pdf>
- Departamento Nacional de Planeacion. (s.f.). *Bases del Plan Nacional de Desarrollo 2014-2018*. Obtenido de <https://colaboracion.dnp.gov.co/cdt/prensa/bases%20plan%20nacional%20de%20desarrollo%2014-2018.pdf>
- DiazGranados Ordóñez, G. F. (2016). *La comercializacion de la Big Data*. Obtenido de Ciencias Juridicas Javeriana: <https://cienciasjuridicas.javeriana.edu.co/documents/3722972/7912168/7-%20LA+COMERCIALIZACI%C3%93N.pdf/17c8d327-d25b-474c-a9c1-1797114e3c0b>
- Ministerio de Justicia y del Derecho. (s.f.). *Estructura Organizacional de la Entidad*. Obtenido de <https://www.minjusticia.gov.co/Ministerio/Estructura-organizacional-de-la-entidad>
- Ministerio de Tecnologías de la Información y las Comunicaciones. (s.f.). *Sistemas de Gestión de la Seguridad de la Información (SGSI)*. Obtenido de <https://www.mintic.gov.co/gestion-ti/Seguridad-TI/Modelo-de-Seguridad/5482:Sistemas-%20de-Gesti%C3%B3n-de-la-Seguridad->

de-la-Informaci%C3%B3n-SGSI

MyABCM. (s.f.). *¿Cómo puede su empresa ser innovada por la cultura basada en datos?* Obtenido de

MyABCM: <https://www.myabcm.com/es/blog-post/cultura-basada-en-datos/>

Valero, M. (2016). *¿Que es una ETL?* Obtenido de Bigeek: <https://blog.bi-geek.com/que-es-una-etl/>

Vega , R., Szyszlican, M., & Borregón, E. M. (s.f.). *Manual Kibana OCDS*. Obtenido de

https://manualkibanaocds.readthedocs.io/es/latest/C2/Seccion1/1_ElasticSearch.html

Villarrubia, C. (2018). *Colombia define una política pública para big data*. Obtenido de

Datacenterdynamics: <https://www.datacenterdynamics.com/es/noticias/colombia-define-una-pol%C3%ADtica-p%C3%BAblica-para-big-data/>