

**PLAN DE GESTIÓN DE CALIDAD DE DATOS PARA MEJORAR LA
OPORTUNIDAD Y PERTINENCIA DE LA INFORMACIÓN DE LA OFERTA
INSTITUCIONAL EN LA DIRECCIÓN DE APROPIACIÓN DEL MINISTERIO TIC**

MAGALY RINCÓN RODRÍGUEZ

FABIO HERNAN OSORIO VILLADA

Asesor

UNIVERSIDAD EXTERNADO DE COLOMBIA

FACULTAD DE ADMINISTRACIÓN DE EMPRESAS

MAESTRÍA EN GERENCIA ESTRATÉGICA DE TECNOLOGÍAS DE LA INFORMACIÓN

BOGOTÁ D.C.

2019

Tabla de Contenido

Tabla de Contenido	2
Listado de Figuras	5
Listado de Tablas.....	6
Listado de Anexos	7
Listado de Abreviaturas	8
Glosario	9
Introducción.....	11
1 Planteamiento del problema	13
1.1 Antecedentes del problema	13
1.2 Problemática.....	15
1.3 Preguntas de investigación.....	16
1.3.1 Principal	16
1.3.2 Secundarias	16
1.4 Objetivos.....	17
1.4.1 Objetivo General.....	17
1.4.2 Objetivos Específicos	17
1.5 Alcance y limitaciones.	18
1.6 Justificación	18
2 Marco Referencial.....	20
2.1 Marco contextual.....	20
2.1.1 Estructura del sector de tecnologías de la información y las comunicaciones	20
2.1.2 Naturaleza Jurídica de MinTIC	21
2.1.3 Misión	21
2.1.4 Visión.....	22
2.1.5 Valores	22
2.1.6 Organigrama	23
2.1.7 Dirección de Apropiación de TI.....	26
2.1.8 Oficina de Tecnologías de la Información	28
2.2 Marco normativo	29
2.2.1 Normatividad Colombiana decreto 1008 de 2018 política de gobierno digital	30

2.2.2	Marco de referencia de arquitectura TI.....	31
2.2.3	Ley 1581 del 2012 Protección de datos personales	33
2.2.4	Ley 1712 del 2014 Transparencia	33
2.2.5	Directiva Presidencial de 2019.....	34
2.2.6	CONPES 3920 de 2018 Explotación de Datos	34
2.2.7	Transformación Digital DNP.....	35
2.3	Marco conceptual y revisión de literatura	36
2.3.1	Calidad de datos.....	36
2.3.2	Gestión de calidad de datos.....	37
2.3.3	Gobierno de datos	38
2.3.4	Referencias de literatura sobre calidad de datos	39
2.3.5	Marcos de referencia para la gestión de calidad de datos.....	41
2.3.6	DAMA.....	41
2.3.7	ISO 8000	45
2.3.8	MIKE 2.0	47
2.4	Buenas prácticas en proceso de calidad de datos	51
2.4.1	Perfilamiento y evaluación de datos.....	51
3	Metodología.....	54
3.1	Diseño de la investigación	54
3.1.1	Selección de la muestra	55
3.1.2	Herramientas para recolección de datos.....	58
3.1.3	Análisis	59
3.2	Paso a paso de la metodología	64
3.3	Fuentes de información: personas, instrumentos.....	67
4	Resultados de la investigación.....	72
4.1	Diagnóstico situacional en la Dirección de Apropiación, respecto a la gestión de datos	72
4.2	Criterios relevantes y buenas prácticas de los marcos de referencia a aplicar.....	82
4.3	Acciones de mejora propuestas a los procedimientos actuales.....	85
4.4	Identificación de los beneficios de un plan de gestión de calidad de datos	87
5	Diseño y propuesta de implementación del plan de gestión de calidad.....	89
5.1	Objetivo	89
5.2	Alcance.....	89

5.3 Estrategia de gobierno de datos.....	90
5.3.1 Impulsores de negocio	91
5.3.2 Misión	91
5.3.3 Visión.....	91
5.4 Objetivos, principios y políticas de gobierno de datos.....	92
5.4.1 Objetivos de gobierno de datos.....	92
5.4.2 Principios de datos.....	93
5.4.3 Políticas de datos	94
5.5 Modelo operacional del plan de Gestión de Calidad de Datos	96
5.5.1 Identificación del macroproceso de la Dirección de Apropiación	98
5.5.2 Procesos de gestión de datos en la Dirección de Apropiación	98
5.5.3 Flujo de la información en la Dirección de Apropiación	102
5.5.4 Procedimientos propuestos para la Dirección de Apropiación	103
5.5.5 Medición del plan de calidad en el marco de gobierno de datos.....	106
5.5.6 Roles y responsabilidades	108
5.5.7 Tecnología.....	109
5.5.8 Plan de implementación	111
5.6 Recomendaciones.....	111
6 Conclusiones	113
Referencias Bibliográficas	115
Anexos.....	121

Listado de Figuras

Figura 1 Ilustración Estructura organizacional del MinTIC / Despacho del Ministro.....	23
Figura 2 Ilustración Estructura organizacional del MinTIC / Secretaría General	23
Figura 3 Ilustración Estructura organizacional del MinTIC / Despacho del Viceministerio de Conectividad y Digitalización	24
Figura 4 Ilustración Estructura organizacional del MinTIC / Despacho del Viceministerio de Economía Digital	25
Figura 5 Ilustración Estructura organizacional del Ministerio de Tecnologías de la Información y Comunicaciones.....	25
Figura 6 Estructura organizacional del Ministerio de Tecnologías de la Información y Comunicaciones, destacando a la Dirección de Apropiación	27
Figura 7 Estructura organizacional adaptada del Ministerio de Tecnologías de la Información y Comunicaciones	29
Figura 8 Momentos de la Política de Gobierno Digital	30
Figura 9 Diagrama conceptual de la política de Gobierno Digital.	31
Figura 10 Cuadro de los seis dominios del Marco de Referencia de AE.....	32
Figura 11 Diagrama de efectos económicos y sociales.....	36
Figura 12 Marco para la Gestión de Datos del DAMA	42
Figura 13 Mapa general de la norma ISO 8000.....	46
Figura 14 Familia de estándares de la ISO 8000	47
Figura 15 Marco de la metodología Mike 2.0	48
Figura 16 Fases de Implementación de la metodología Mike 2.0	49
Figura 17 Enfoque para medir la Calidad de Datos y ejemplo de medición.....	50
Figura 18 Diagrama de pasos para la evaluación de calidad de datos	53
Figura 19 Fórmula de poblaciones finitas para el tamaño de la muestra 1.....	55
Figura 20 Fórmula de poblaciones finitas para el tamaño de la muestra 2.....	57
Figura 21 Fórmula de poblaciones finitas para el tamaño de la muestra 3.....	57
Figura 22 Modelo de gobierno de datos a alto nivel	97
Figura 23 Mapa de Macroprocesos del MinTIC	98
Figura 24 Diagrama de flujos de los procesos formales y tácitos de Gestión de Datos en la Dirección de Apropiación	99
Figura 25 Ciclo del proceso de Calidad de los Datos.....	101
Figura 26 Flujo de la información en la Dirección de Apropiación	103
Figura 27 Métricas propuestas para medir la Calidad de Datos.....	108
Figura 28 Matriz de Roles y Responsabilidades RACI	109

Listado de Tablas

Tabla 1 Elementos de calidad de datos.	43
Tabla 2 Organización de transcripción de entrevistas por parrillas segmentadas	61
<i>Tabla 3 Concentración de ideas por bloque temático.....</i>	<i>62</i>
Tabla 4 Triangulación Bloque Temático 2 Y MIG	63
Tabla 5 Triangulación bloque temático 2 y respuestas de expertos.	64
Tabla 6 Resumen de la Metodología	67
Tabla 7 Criterios teóricos de selección de la muestra 1	68
Tabla 8 Criterios teóricos de Selección de la Muestra 3	69
Tabla 9 Tabla de preguntas orientadoras	71
Tabla 10 Perfilamiento de longitud de campos	73
Tabla 11 Perfilamiento de campos nulos.....	74
Tabla 12 Perfilamiento de tipo de datos y formato.....	75
Tabla 13 Matriz de diagnóstico aplicada.....	75
Tabla 14 Campos con problemas de exactitud	76
Tabla 15 Campos con problemas de consistencia	77
Tabla 16 Campos con problemas de Validez	78
Tabla 17 Resultados de la evaluación diagnóstica.....	79
Tabla 18 Estructura propuesta de una Política de Gestión de Datos	95
Tabla 19 Tabla de procedimientos propuestos.....	106
Tabla 20 Tabla de indicadores de medición propuestos	107
Tabla 21 Tabla de características de la Open Refine	110

Listado de Anexos

Anexo 1 Guía I de entrevista a responsables de gestión de datos.....	121
Anexo 2 Guía II de entrevista a expertos en gestión de datos.	124

Listado de Abreviaturas

ANE: Agencia Nacional del Espectro.

CIO: Chief Information Officer.

CRC: Comisión de Regulación de Comunicaciones.

CRUD: Create, Read, Update, Delete, en español, crear, leer, actualizar y eliminar.

DAMA: Data Management Association. Entidad de Conocimiento para la Gestión de Datos.

DANE: Departamento Administrativo Nacional de Estadística.

DNP: Departamento Nacional de Planeación.

eBook: electronic book. Libro Electrónico.

ETL: Extract, Transform and Load. Extracción, Transformación y Carga. Elemento importante que permite la integración de sistemas y aplicaciones.

IBM: International Business Machines. Multinacional de tecnología informática y consultoría.

IoT: Internet of Things o Internet de las cosas.

MinTIC: Ministerio de Tecnologías de la Información y las Comunicaciones.

MIG: Modelo Integrado de Gestión.

MIT: Massachusetts Institute of Technology.

OCDE: Organización para la Cooperación y el Desarrollo Económico.

RTVC: Radio Televisión de Colombia.

TI: Tecnologías de Información.

TIC: Tecnologías de la Información y las Comunicaciones.

Glosario

Activo: son los bienes tangibles o intangibles que posee una organización y que puede convertirse en dinero corriente si éste es vendido.

Área Misional: es un área que se encarga de realizar las actividades contempladas en la cadena de valor de una organización.

Analítica: herramientas y técnicas que se usan para el análisis de datos a fin de explorar patrones y tendencias.

Arquitectura: se define así, a las actividades de idear, diseñar y construir esquemas base para la implementación de alguna solución.

Big Data: término que se utiliza para describir el gran volumen de datos tanto estructurados como tablas, bases de datos, y no estructurados como texto, video e imágenes.

Brecha Digital: se le denomina así, a la desigualdad que existe entre personas en cuánto al acceso, uso y apropiación de las tecnologías.

Core: Término que se utiliza para destacar alguna característica como principal o primordial en un contexto determinado.

Cultura Digital: concepto describe toda actividad relacionada o influenciada por las tecnologías de información.

Dimensiones: se denomina así, a las características o atributos que describen mejor a los datos.

Estándar: es la unificación de características, reglas y guías propuestas por un cuerpo de conocimiento reconocido y que son aceptadas por la mayoría en un medio o sector determinado para un fin específico.

Inteligencia de Negocios: es un conjunto de procesos, herramientas y tecnología que permiten llevar a cabo la transformación de los datos en información y esta información en conocimiento útil para la toma de decisiones.

Interoperabilidad: es la capacidad que presentan dos o más sistemas para establecer comunicación para intercambio e interpretación de información.

Lineamiento: es una orientación o directriz que debe ser acatada por el entorno que éste lo defina.

Metodología: conjunto de métodos organizados de manera racional para cumplir con un objetivo.

MIG: es el modelo integrado de gestión del Ministerio TIC, en donde se encuentran todos los procesos, procedimientos, manuales, guías sobre las cuales se rige la Entidad.

Oferta Institucional: hace referencia a la oferta de planes, programas, proyectos e iniciativas que tiene a su disposición el Ministerio de Tecnologías para todos los ciudadanos.

Operadores de Servicios: hace referencia a todas las empresas que contratan con el Ministerio para la prestación de algún bien o servicio, que permita cumplir con algunos objetivos misionales, y que producto de dicha ejecución, captura y recopila datos que luego son entregados al Ministerio como parte de los resultados.

Política: es la centralización de una idea para ejercer poder sobre un grupo determinado de personas.

Procedimiento: un conjunto de pasos a seguir de forma sistemática para alcanzar un objetivo.

Requisitos: son las necesidades que se identifican y definen para satisfacer una condición determinada.

Usuarios: personas responsables de la administración y gestión de los datos en cada uno de sus proyectos.

Web Service: es un mecanismo de comunicación entre máquinas conectadas que usan una serie de protocolos y estándares para intercambiar datos.

Introducción

La Transformación Digital, es probablemente uno de los términos de la actualidad más usado en las diferentes industrias del país; es por ello, que el Gobierno Nacional viene desarrollando políticas que fortalecen este proceso en las organizaciones tanto públicas como privadas, así como lo manifiesta el Departamento Nacional de Planeación [DNP] quien "...está trabajando en la Política Nacional para la Transformación Digital con el fin de aumentar la generación de valor social y económico a través del uso estratégico de tecnologías digitales" (2019, p.p. 4).

Es en éste entorno donde "lo digital convierte los procesos en datos" (PowerData, 2018, p.p. 17) en donde se ha producido un crecimiento vertiginoso de los mismos a través del aprovechamiento de nuevas tecnologías, permitiendo que las organizaciones puedan ver sus objetivos a través de los datos; razón por la cual se convierte en un activo importante para las organizaciones, considerándose como uno de los apalancadores de la Transformación Digital junto con la Estrategia y la Tecnología (Gati Consultores, 2019). Es así que el presente trabajo se enfoca en el apalancador; *Datos*.

Acorde con lo anterior, el presente trabajo aborda como primer capítulo la problemática actual de la Dirección de Apropiación del MinTIC sobre la calidad de sus datos, partiendo de la posición del autor como parte de la Oficina de TI a la cual pertenece, quien evidencia la problemática y de quien surge la idea de una propuesta de solución entendiendo la importancia de una buena gestión de datos al interior de la entidad como una necesidad primordial tal como lo señala Peters (2001) citado por DAMA (2008) "las organizaciones que no entienden la gran importancia de la gestión de datos y la información como activos tangibles en la nueva economía no van a sobrevivir" (p. 11).

El capítulo dos, parte de la revisión documental acerca de la composición y contexto de la organización sobre la que se desarrolla el proyecto. También se encuentra el marco legal que le compete con relación a los datos, resaltando normatividades relevantes como: las leyes 1581 y 1712, CONPES de explotación de datos y de transformación digital y el Marco de Referencia de la Arquitectura Empresarial del decreto 1008 de Gobierno Digital. Este capítulo contiene además la revisión de literatura descrita en el marco conceptual la cual se enmarca en los objetivos del proyecto, en el cual se presentan referentes conocidos a nivel internacional en temas de gestión de datos como el DAMA, los cuales son tenidos en cuenta para la posterior construcción del plan de gestión de calidad.

Siguiendo el curso de su desarrollo, en el capítulo tres se describe el diseño metodológico empleado en la investigación definiendo diferentes etapas que

permiten; la selección de los dos tipos de muestras que se identificaron (personas, y bases de datos), la definición de los instrumentos de recolección de datos tales como entrevistas enfocadas a los responsables de la gestión de datos y a expertos, así como la matriz de diagnóstico para la evaluación de una de las bases de datos seleccionadas bajo muestreo teórico, también se aborda la forma como se realizó el análisis de los instrumentos aplicados.

En el capítulo cuatro, se realiza la descripción del análisis realizado luego de la aplicación de los instrumentos diseñados, tales como las guías de entrevista semiestructurada, en las cuales se logró identificar que hay un problema latente de calidad en los datos desde las fuentes y que entre las principales causas identificadas están: la falta de lineamientos para las exigencias mínimas a los operadores de servicio en los procesos de captura de datos, la falta de controles de calidad en los procesos de revisión por parte de los responsables de la gestión de datos, la falta de procedimientos formales establecidos para la gestión de datos y por último la falta de una cultura organizacional entorno a la gestión de datos.

Por último, el capítulo cinco concluye con la elaboración de un plan de gestión de calidad de datos sugerido para optimizar la gestión de los procesos en la Dirección de Apropiación del Ministerio TIC, concebido como una herramienta que permite mejorar las características de los datos generando un incremento en la calidad de los resultados obtenidos en el proceso de reporte de información, brindando de este modo mejores oportunidades para la toma de decisiones.

1 Planteamiento del problema

Los datos se definen como hechos o cosas conocidas sobre los cuales se infiere para posteriormente producir la información que se requiere analizar (Birnbaum, 2007a). Los datos que se perciben por los sentidos humanos, de manera continua y sin interrupciones, se denominan analógicos, cuando se interpretan a través de codificación binaria, se denominan datos digitales (Consejo Nacional de Política Económica y Social & Departamento Nacional de Planeación [CONPES], 2018).

La información por su parte es el resultado de los procesos de recolección, almacenamiento y procesamiento de los datos. Cuando se realiza una gestión adecuada de la información, esta logra producir nuevo conocimiento que resulta de gran utilidad para las organizaciones (Birnbaum, 2007b) y que da paso a la generación de valor. Cuando los datos son procesados mediante capacidades tecnológicas y son analizados con un fin, se produce información con valor agregado, lo que en las entidades de gobierno se traduce en generar valor social y económico al lograr tomar mejores decisiones y mejorar procesos que repercutan en una mejor calidad de vida para la sociedad.

En el presente capítulo, se hará una descripción de la problemática que determina la necesidad de un plan de gestión de calidad para los datos resultantes de la ejecución de los proyectos de la oferta Institucional, en la Dirección de Apropiación de TI en adelante (Dirección de Apropiación) del Ministerio de Tecnologías de Información y las Comunicaciones en adelante (MinTIC).

1.1 Antecedentes del problema

En la actualidad, vivimos en una era digital en donde el desarrollo de la tecnología de la información ha permitido que se produzca una gran cantidad de datos que se transforman en información de valor, para crear desde oportunidades de negocio, hasta promover la toma de decisiones en un Gobierno. Es así como la información se viene convirtiendo en el activo más importante de las empresas (Westerman, Bonnet, & McAfee, 2014).

La importancia de los datos en la actualidad a nivel mundial, se hace cada vez más evidente, así, por ejemplo, según Mar (2018) en un artículo de la revista Forbes señala que se producen 2,5 quintillones de bytes de datos cada día. Cifra que se considera en aumento por el crecimiento explosivo de los datos promovidos

por tendencias como el Internet de las Cosas (IoT); sólo en los dos últimos años se generó el 90% de los datos en el mundo. Por mencionar algunas cifras de los medios que más datos producen cada minuto según infografía anual de DOMO (2019) encontramos: 511.200 mensajes en Twitter, 4.497.420 búsquedas son realizadas en Google, 4.500.000 vídeos son vistos en YouTube, Skype registra 231.840 llamadas, los usuarios de Instagram comparten 55.140 imágenes, tan sólo en Estados Unidos se usan 4.416.720 GB de datos de Internet (Gigabyte es una unidad de medida aproximadamente igual a 1 billón de bytes), se procesan 162.037 transacciones a través de Venmo y, por si fuera poco, alrededor de 188 millones de correos electrónicos son enviados x.

Dado lo anterior y gracias al volumen de datos y al descubrimiento de su valor, se ha venido concentrando una mayor atención sobre éstos y se hacen contundentes expresiones, tales como la del responsable de Global Client Solutions del BBVA en un artículo quien dice que "...los datos serán la clave del éxito en el futuro de las empresas" (White, 2019, p.1).

Por otra parte, la preocupación por la calidad de los datos va en aumento, gracias al surgimiento de procesos involucrados en las nuevas tecnologías de la información tales como, la inteligencia de negocios, el big data y la analítica que permiten su explotación y aprovechamiento, generando valor en las diferentes industrias como en el campo médico, logística y las telecomunicaciones (Cai & Zhu, 2015b), por lo que "...es necesario la obtención de datos precisos y de alta calidad" (Cai & Zhu, 2015c). Es por ello, que debido a los bajos niveles de calidad que presentan las bases de datos, algunas empresas se ven obligadas en ocasiones a limitar el alcance de sus iniciativas analíticas y a su vez, incurrir en grandes costos recursos tanto financieros, como de tiempo por procesos de detección y corrección de errores (PowerData, s.f.).

Ahora bien, desde la Oficina de TI y como autor del presente trabajo de grado en el marco de los procesos de gestión de información, se ha podido detectar que: en la Dirección de Apropiación del MinTIC, quien tiene bajo su responsabilidad los programas de uso y apropiación masivo de las tecnologías de información, las fuentes de datos presentan inconsistencias. Por lo tanto, las acciones que se tomen frente a este problema son determinantes para el mejoramiento de los reportes de información a generar y así mismo para el aumento de la eficiencia en tiempos invertidos en este proceso.

1.2 Problemática

Todas las organizaciones a nivel mundial sean privadas o públicas, se encuentran generando diariamente información de diferentes tipos y desde diferentes fuentes tales como: conjuntos de datos de internet, datos de internet de las cosas y datos recopilados por diferentes industrias (Cai & Zhu, 2015d). Dichas organizaciones, cuando presentan necesidades de análisis de información que les sirva de soporte para la toma de decisiones basada en datos, inician proyectos de analítica como inteligencia de negocios, Big Data, inteligencia artificial, entre otros, los cuales dice (SAS Latin America, 2016a) en su etapa inicial no logran detectar todos los problemas asociados a la gestión de los datos.

Muestra de ello ha sido el MinTIC, que con el incremento de solicitudes de información que provienen de diferentes actores, no ha sido ajeno a la adopción de tendencias y soluciones tecnológicas que facilite el proceso de gestión de información en la entidad. No obstante, en el marco de esta estrategia, los procesos para la explotación de datos y la toma de decisiones se han visto afectados al enfrentarse con la no conformidad de los datos que registra la entidad, los cuales son entregados en su mayoría por agentes externos como proveedores y/o operadores de servicios (empresas que contratan con el Ministerio la prestación de algún bien o servicio) que ejecutan los proyectos de la oferta institucional de la Dirección de Apropiación; entre otras cosas, las faltas más comunes que se han venido detectando en los datos, son : oportunidad, validez, integridad y exactitud.

Es así como, para lograr reportes confiables, las áreas se ven sometidas a una serie de reprocesos manuales que conllevan a disponer de tiempos considerables por parte de las personas responsables del manejo de los datos, a fin de poder entregar información a la alta gerencia. Por consiguiente, genera impactos que representan costos en términos de tiempo y oportunidad por las ineficiencias causadas. Estos costos en cálculos estimados por el autor del presente trabajo significan para el caso de la Dirección de Apropiación, que reporta información de 5 proyectos de la oferta institucional, poner a disposición un recurso humano por 8 horas diarias durante los 5 primeros días del mes, para cada proyecto. Esto significa que para la Dirección representa 200 horas (5 proyectos x 5 días/proyecto x 8 horas /día) equivalentes a \$10.351.400 (diez millones trescientos cincuenta y un mil cuatrocientos pesos) mensuales enfocados solo a la ejecución de dicha actividad, considerando el costo por hora acorde con el tope máximo de 15 SMMLV para la prestación de servicios profesionales de apoyo a la gestión según ([MinTIC], 2017) a \$51.757 (cincuenta y un mil setecientos cincuenta y siete pesos).

Con todo lo anterior y con el paso del tiempo, se han evidenciado algunas consecuencias en las que las áreas empiezan a visualizar que los nuevos datos no se ajustan a los anteriores, se detectan problemas de rendimiento en el procesamiento de los datos y por lo tanto se genera falta de credibilidad y oportunidad en la información entregada. Estos problemas de rendimiento se ven reflejados en la ejecución de los procesos de extracción, transformación y carga cada mes, en adelante *ETL* (Extract, Transform and Load); lo que se espera es que se puedan ejecutar solo una vez y de manera automática dentro de los tiempos que se tienen programados; sin embargo, desde la Oficina de TI se han realizado en varias ocasiones más de 3 iteraciones con los responsables del manejo de los datos de la Dirección, hasta que éstos ya han estado conformes y han podido ser procesados finalmente por los sistemas.

1.3 Preguntas de investigación

1.3.1 Principal

¿Cuál es la mejor forma para que la Dirección de Apropiación pueda ejercer control sobre la calidad, homologación y estandarización de los datos capturados por los operadores de servicio de la oferta institucional, a fin de mejorar la pertinencia y oportunidad de la información?

1.3.2 Secundarias

- ¿Cuál es el diagnóstico de calidad de datos y las principales causas de la problemática que se presenta en la Dirección de Apropiación relacionada con la gestión de datos?
- ¿Cuál podría ser el marco de referencia que se permite adaptar y adoptar de mejor forma, para atender las necesidades de gestión de datos de la oferta institucional en la Dirección de Apropiación?
- ¿Cuál sería la propuesta de mejora a los procedimientos actuales de gestión de datos en la Dirección de Apropiación?
- ¿Cuáles podrían ser los beneficios de proponer un plan de gestión de calidad de datos para la Dirección de Apropiación?

- ¿Cómo podría ser el plan de gestión de calidad de datos que permita en el corto, mediano y largo plazo cerrar brechas y obtener mejoras en la oferta institucional de la Dirección de Apropiación?

1.4 Objetivos

1.4.1 Objetivo General

Definir un plan para la gestión de calidad de datos adaptado a la Dirección de Apropiación, que se base en las buenas prácticas de los marcos de referencia mencionados en el presente trabajo, para ejercer control sobre la calidad, homologación y estandarización de los datos capturados por los operadores de servicio de la oferta institucional, a fin de mejorar la pertinencia y oportunidad de la información.

1.4.2 Objetivos Específicos

- Realizar un diagnóstico de calidad de datos e identificar las principales causas de la problemática que describan la situación actual en la Dirección de Apropiación, respecto a la gestión de datos.
- Evaluar criterios relevantes y seleccionar buenas prácticas de los marcos de referencia mencionados en el presente trabajo, para adaptarlos en un plan que atienda las necesidades de gestión de datos de la oferta institucional en la Dirección de Apropiación.
- Proponer las acciones de mejora a los procedimientos actuales de gestión de datos e identificar los beneficios de proponer un plan de gestión de calidad de datos para la Dirección de Apropiación
- Proponer un plan de gestión de calidad basado en las acciones de mejora enlistadas y los criterios seleccionados de los marcos de referencia, que permita en el corto, mediano y largo plazo cerrar brechas y obtener mejoras en los datos de la oferta institucional de la Dirección de Apropiación.

1.5 Alcance y limitaciones.

El alcance de este trabajo se enfoca en proponer un plan de gestión de calidad de datos, en la Dirección de Apropiación del Ministerio de Tecnologías de Información y las Comunicaciones para los cinco (5) proyectos de la oferta institucional de esta Dirección; Teletrabajo, Ciudadanía Digital, EnTicConfío, ConverTIC y Redvolución los cuales actualmente son los proyectos que reportan los datos resultantes de su ejecución a los interesados.

Para la investigación no se tiene en cuenta de manera directa a los Operadores de Servicios, toda vez que la información relacionada con los mismos es suministrada por los responsables de la gestión de datos de la Dirección, ya que son ellos los conocedores directos de la problemática que viven y de la gestión directa con estos actores.

Por otra parte, se tiene como limitación en el marco de las unidades muestrales, la accesibilidad que tenga el investigador a éstas.

Dicho plan es propuesto usando las buenas prácticas seleccionadas en el desarrollo del presente trabajo. No incluye su implementación dado que es potestad del MinTIC su adopción o no.

El autor del presente proyecto de grado trabaja en el MinTIC en la Oficina de Tecnologías de la Información, encontrando desde el dominio de información del marco de referencia de arquitectura empresarial algunas debilidades en la gestión de datos de la oferta institucional.

El plan está centrado en la gestión de datos soporte a la oferta institucional de la Dirección de Apropiación y no incluye propuestas de nuevas ofertas institucionales o modificación de las actuales.

1.6 Justificación

El diseño de un plan para la gestión de calidad de datos es fundamental para mejorar procesos de reporte, análisis y sobre todo, para que sea posible la toma de decisiones de manera asertiva (Deusto Formación, 2018), ya que Loucks et al. (2016) menciona que “datos malos hacen imposible una buena toma de decisiones” citado por (Velthuis, 2018a, p. 26). Así también, reduce impactos negativos como el reprocesamiento de información, tiempos de las personas que realizan dicha labor, pérdida de credibilidad y altos costos, que por ejemplo, según estima el Data Waterhousing Institute, sólo en los Estados Unidos significa un total de 600.000 millones de dólares por año (Méndez del Río, 2006).

Dicho plan puede ser ejecutado en las áreas de la entidad y puede extenderse su aplicación a nivel del Sector TIC; sin embargo, el presente proyecto

se enfoca en un área del MinTIC, ubicada dentro del Viceministerio de Economía Digital; la Dirección de Apropiación, en donde se hace necesario obtener los beneficios de lo que encierra el concepto de calidad de datos según SAS Latin America (2016b) “en donde lo ideal es tener menor retrabajo y sí mayor confianza en los datos” (p.1).

A consecuencia de la creciente demanda de información en la entidad para la generación de reportes que soportan la toma de decisiones, se hace notable la necesidad de un plan de gestión de datos, que le permita a la Dirección de Apropiación no sólo tomar consciencia de su importancia, sino también auto gestionarse y empoderarse de esta herramienta a través de su aplicación, de manera que los datos que se reciban de los operadores de servicio que ejecutan los proyectos, ofrezcan a la Dirección de Apropiación una óptima calidad en los datos para obtener provecho de éstos y generar valor. Así como menciona Ramírez (2017) “se requiere un cambio profundo en la mentalidad empresarial, para convertir el dato en el principal activo de la compañía y su explotación en la mayor prioridad.” Citado por (Velthuis, 2018b, p. 27).

Es por ello, que el presente trabajo propone un plan de gestión, que le permita a los responsables del manejo de datos de la oferta institucional en la Dirección de Apropiación, medir y revisar la calidad de éstos, a fin de poder exigir a terceros los mínimos aspectos en la calidad de los datos requeridos. Por lo tanto, la implementación del plan de gestión mejoraría notablemente el comportamiento de los datos en los reportes, generando mayor credibilidad, al mismo tiempo que facilitaría las labores diarias de los funcionarios y/o servidores de la entidad, reduciendo la ejecución de tareas como corrección y reprocesamiento de bases de datos (Corporación Colombia Digital [CCD], 2017).

Atendiendo a las consideraciones anteriores, es necesario que el MinTIC como ente encargado de las Tecnologías de la Información y las Comunicaciones y en especial la Dirección de Apropiación, adopte mecanismos relacionados a la adecuada gestión de datos, así como también identifique la notoriedad del problema relacionado que se está presentado, dado que, según Calero, Moraga, & Piattini (2010) es un tema que debe manejarse a nivel organizacional. Por esta razón, se requiere implementar herramientas y controles que sean de utilidad en su aplicación y que generen valor agregado a los procesos de información de la Dirección, sin incurrir en altos costos de soluciones tecnológicas. En este sentido, un plan de gestión atiende a dicha necesidad ya que puede ser usado como buena práctica y se convierte en una medida preventiva de gran importancia.

2 Marco Referencial

En el presente capítulo se presenta un análisis del contexto organizacional, dando a conocer cómo desde la Oficina de TI, se apoyan los procesos de gestión de información en la entidad, también se presentan las herramientas y metodologías disponibles para construir una solución adecuada según las necesidades de la Dirección de Apropiación en calidad de datos.

2.1 Marco contextual

El Ministerio de Tecnologías de Información y las Comunicaciones es consciente de que la misión de las instituciones es servir a los ciudadanos con un Estado abierto y transparente; con servicios y trámites ágiles y efectivos, con información precisa y de alta calidad (Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC], 2019). La oficina de TI desde la dimensión de información está en la búsqueda permanente de mecanismos y/o herramientas para aportar al desarrollo de la entidad mediante la definición de lineamientos que les permita a las áreas, ser implementadas para un mejor desempeño de sus actividades en los ámbitos de información.

El plan de gestión que se propone en este trabajo está adaptado a las necesidades específicas de la Dirección de Apropiación en temas de calidad de datos y que son diseñados desde la oficina de TI, por lo tanto, es necesario tener claro el contexto de la entidad para facilitar un mejor entendimiento del propósito del proyecto.

2.1.1 Estructura del sector de tecnologías de la información y las comunicaciones

De acuerdo con el decreto único del Sector TIC en el artículo 1.1.1.1 se asigna como cabeza de sector al Ministerio de Tecnologías de Información y las Comunicaciones, para que ejerza sus funciones como lo define la ley 1341 de 2009.

Dentro de este sector, existen también entidades adscritas al Ministerio que cumplen funciones de soporte y mejoramiento de procesos en el desarrollo de infraestructura, servicios y aplicaciones para beneficiar a usuarios y entidades sectoriales que aportan al crecimiento del país (Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC], 2014).

A continuación, se listan las entidades adscritas sobre las cuales el Ministerio actúa como cabeza de sector:

- Fondo Único de Tecnologías de la Información y las Comunicaciones.
- Agencia Nacional del Espectro (ANE)
- Comisión de Regulación de Comunicaciones (CRC)
- Radio Televisión de Colombia (RTVC)- Señal Colombia (Sistema de Medios Públicos)
- Red Postal de Colombia 4-72.

Después de conocer cuál es la estructura para el sector TIC, se presenta el ámbito de aplicación sobre el cual actuarán estas entidades anteriormente mencionadas, considerando el decreto único del sector en su artículo 2.1.1.2 el cual dice de la siguiente manera:

El presente decreto aplica a los proveedores de redes y servicios de telecomunicaciones, a los proveedores del servicio de radiodifusión sonora, a los operadores de servicios postales, a las personas públicas y privadas que las disposiciones de este decreto determinen y en general, a las entidades del sector de Tecnologías de la Información y las Comunicaciones (MinTIC, 2015 p. 18).

2.1.2 Naturaleza Jurídica de MinTIC

El Ministerio de Tecnologías de la Información y las Comunicaciones, según la Ley 1341 o Ley de TIC, es la entidad que se encarga de diseñar, adoptar y promover las políticas, planes, programas y proyectos del sector TIC (MinTIC, 2019a) .

2.1.3 Misión

Promover el acceso, uso efectivo y apropiación masivos de las TIC, a través de políticas y programas (MinTIC, 2019b); lo anterior, siempre en busca del mejoramiento de la calidad de vida de los colombianos aportando de esta manera al desarrollo sostenible del país.

Dentro de los factores clave de la misión, encontramos el acceso, uso efectivo y apropiación de las TIC, esto conduce a la ejecución de proyectos en todas las regiones del país en conjunto con el sector privado para lograr alcanzar

a los ciudadanos en pro de beneficiarlos, tanto en temas de infraestructura de redes como de sensibilización de personas en el uso de las TIC. Estos proyectos son denominados proyectos misionales o de la Oferta del institucional del MinTIC.

2.1.4 Visión

En el 2018, el Ministerio de las Tecnologías de la Información y las comunicaciones estará preparado para enfrentar los retos de la economía digital y habrá logrado posicionar al país como referente en el desarrollo de aplicaciones sociales y contenidos digitales (MinTIC, 2019a).

Los logros de esta visión para el 2018, se pueden verificar en el Informe al Congreso - 2018 (publicado 19 de julio de 2018) .

2.1.5 Valores

Los valores que como entidad del orden nacional caracterizan al Ministerio TIC son: la honestidad basada en la verdad, el respeto hacia todas las personas sin importar su condición, el compromiso que tienen los servidores públicos en relación con sus funciones, la diligencia para atender las responsabilidades de manera oportuna, la justicia actuando de manera imparcial en cualquier situación y la responsabilidad que atañe sobre las decisiones que se toman (MinTIC, 2019a).

2.1.6 Organigrama

DESPACHO DEL MINISTRO

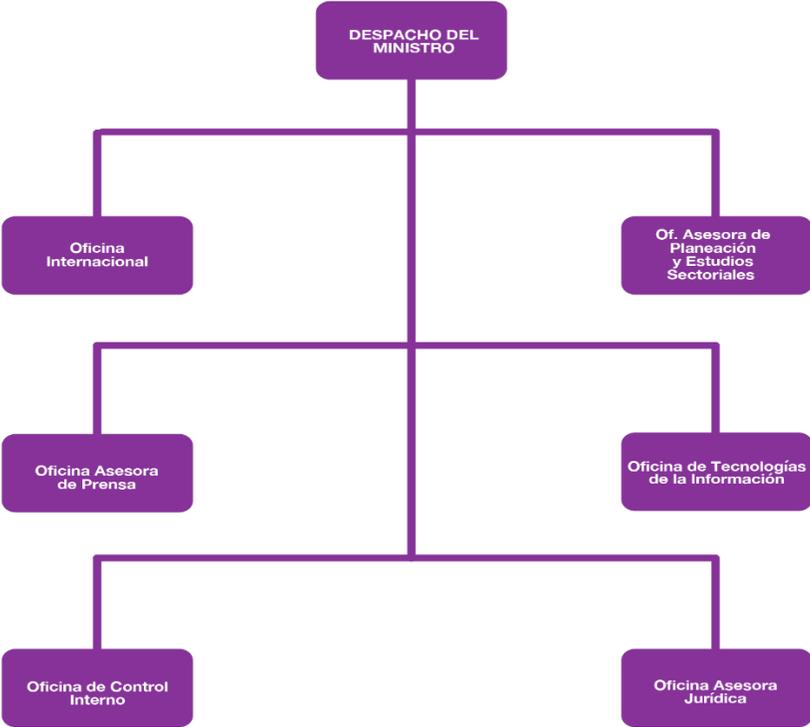


Figura 1 Ilustración Estructura organizacional del MinTIC / Despacho del Ministro.
Fuente:([MinTIC], 2019a)

SECRETARÍA GENERAL



Figura 2 Ilustración Estructura organizacional del MinTIC / Secretaría General
Fuente:([MinTIC], 2019b)

DESPACHO DEL VICEMINISTRO DE CONECTIVIDAD Y DIGITALIZACIÓN

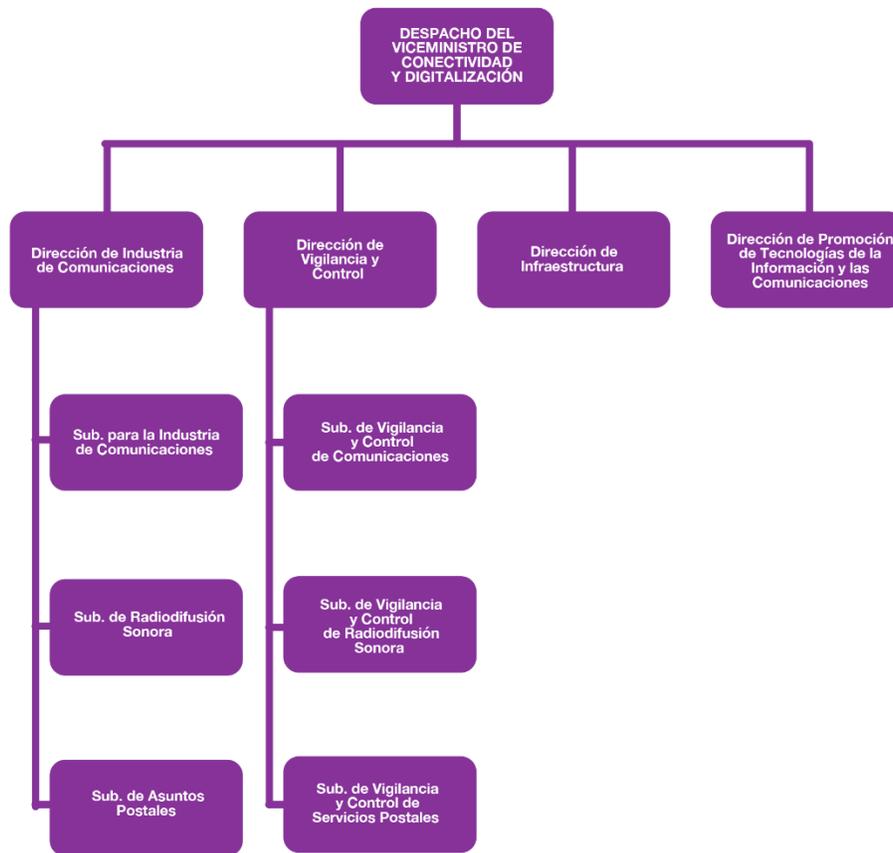


Figura 3 Ilustración Estructura organizacional del MinTIC / Despacho del Viceministerio de Conectividad y Digitalización
Fuente:([MinTIC], 2019c)

DESPACHO DEL VICEMINISTRO DE ECONOMÍA DIGITAL

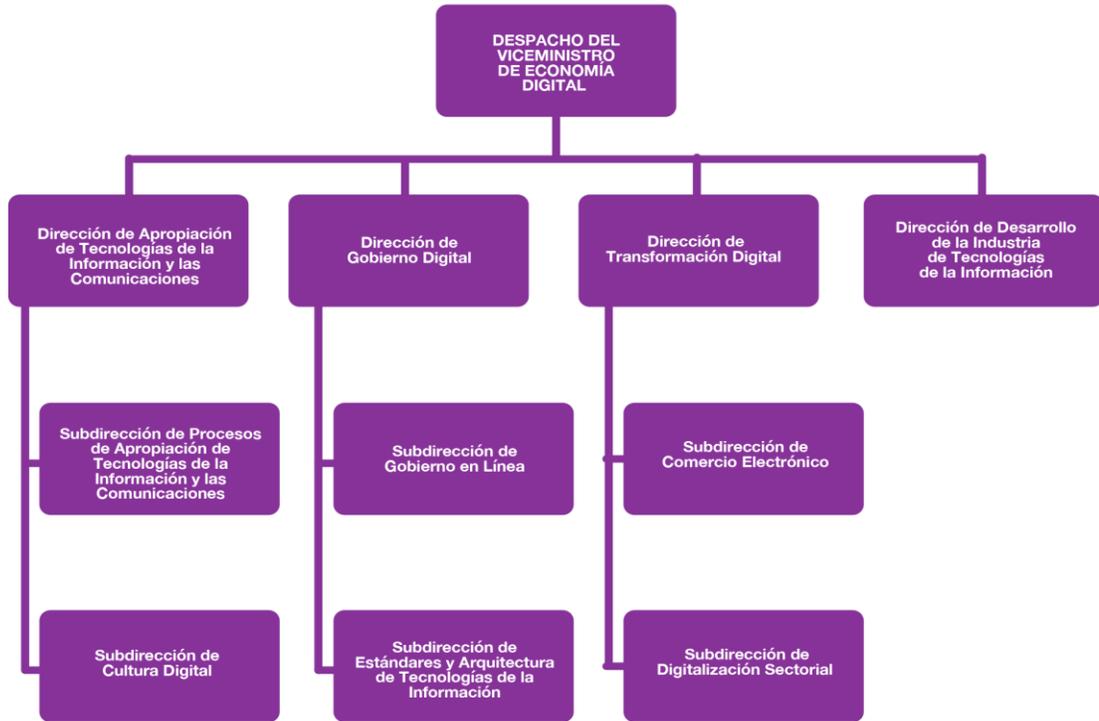


Figura 4 Ilustración Estructura organizacional del MinTIC / Despacho del Viceministerio de Economía Digital
Fuente: ([MinTIC], 2019d)

ÓRGANOS DE ASESORÍA Y COORDINACIÓN



Figura 5 Ilustración Estructura organizacional del Ministerio de Tecnologías de la Información y Comunicaciones.
Fuente:([MinTIC], 2019e)

Se observa en efecto en el organigrama, la manera como están organizadas las diferentes áreas del Ministerio. Lo primero que se puede ver, son las oficinas que dependen directamente del despacho, las cuales son de apoyo estratégico para quien haga sus veces de Ministro(a).

Luego de ellas, viene la secretaría general y sus dependencias, quienes son las encargadas de administrar los recursos financieros, de funcionamiento y humanos para garantizar el buen desempeño de la entidad.

Posterior a ésta, se desglosan los dos viceministerios, el de Conectividad y Digitalización y el más reciente de Economía Digital. Es justo en estos dos viceministerios en donde se generan, promueven y ejecutan todos los proyectos e iniciativas diseñadas desde el plan de gobierno para el cumplimiento de metas y objetivos de cara a los ciudadanos y para regular a los prestadores de servicio que se encuentran en el ámbito de tecnologías de información y comunicaciones, de acuerdo con el decreto único del sector TIC. En estas áreas, es donde se generan los procesos de reportes de información hacia la alta gerencia en medio de su gestionar del día a día por velar por el cumplimiento de los proyectos y las regulaciones que deben realizar según corresponda a cada una.

Estas áreas misionales, mediante la ejecución de proyectos y a través de contratos y/o convenios interadministrativos permiten llevar a cabo los objetivos propuestos; estos contratos generan de vuelta al Ministerio, información que en su gran mayoría es reportada en archivos de Excel y en muy contados casos a través de web services que interoperan con la entidad, por lo que éstas son finalmente las fuentes de información que utilizan las áreas para generar sus reportes y seguimientos.

2.1.7 Dirección de Apropiación de TI

La Dirección de Apropiación de Tecnologías de Información y las Comunicaciones como área de ejemplificación para este proyecto, desarrolla sus programas e iniciativas buscando promover el uso efectivo y la apropiación masiva de las TIC. Este objetivo se logra a partir del diseño y la implementación de proyectos que tienen como meta mejorar la calidad de vida de cada colombiano, reducir la brecha digital, disminuir la pobreza y contribuir al desarrollo del país con la apropiación de las TIC (MinTIC, 2019b).

Se entiende por uso, la capacidad y/o competencia de los Grupos de Interés para utilizar efectivamente las TIC. Por su parte, con el término apropiación nos

referimos a la acción y el resultado de tomar para sí las TIC, generando cambios en la vida cotidiana de los Grupos de Interés del Ministerio (MinTIC, 2019c).

Entre las funciones principales se encuentran: la elaboración de propuestas para la formulación de política pública orientados a desarrollar capacidades que faciliten el uso y apropiación de las Tecnologías, facilitar el acceso a los ciudadanos, diseñar e implementar estrategias de alfabetización digital, articular las demás instancias del Ministerio para lograr el acceso de los ciudadanos a las Tecnologías de Comunicación, a través de procesos de apropiación y Definir y adoptar indicadores básicos de seguimiento y monitoreo a los procesos de generación de competencias básicas para el uso de las Tecnologías de Comunicación en Colombia.

La oferta institucional de esta Dirección busca beneficiar a la población con discapacidad a través del uso de las TIC, para ello, cuenta con programas como Centros de Relevó, Cine para Todos y ConVerTIC. En el ámbito de cultura digital, se enfoca a promover el uso adecuado de internet, redes sociales y la participación ciudadana mediante las iniciativas En TIC confío, Ciudadanía Digital, Teletrabajo y Escuela TIC familia.

DESPACHO DEL VICEMINISTRO DE ECONOMÍA DIGITAL

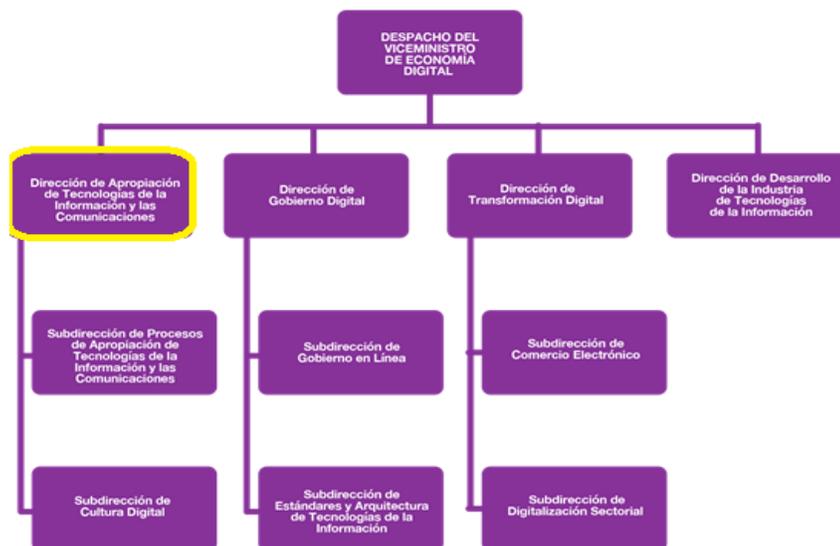


Figura 6 Estructura organizacional del Ministerio de Tecnologías de la Información y Comunicaciones, destacando a la Dirección de Apropiación

Fuente: ([MinTIC], 2019f)

2.1.8 Oficina de Tecnologías de la Información

DESPACHO DEL MINISTRO

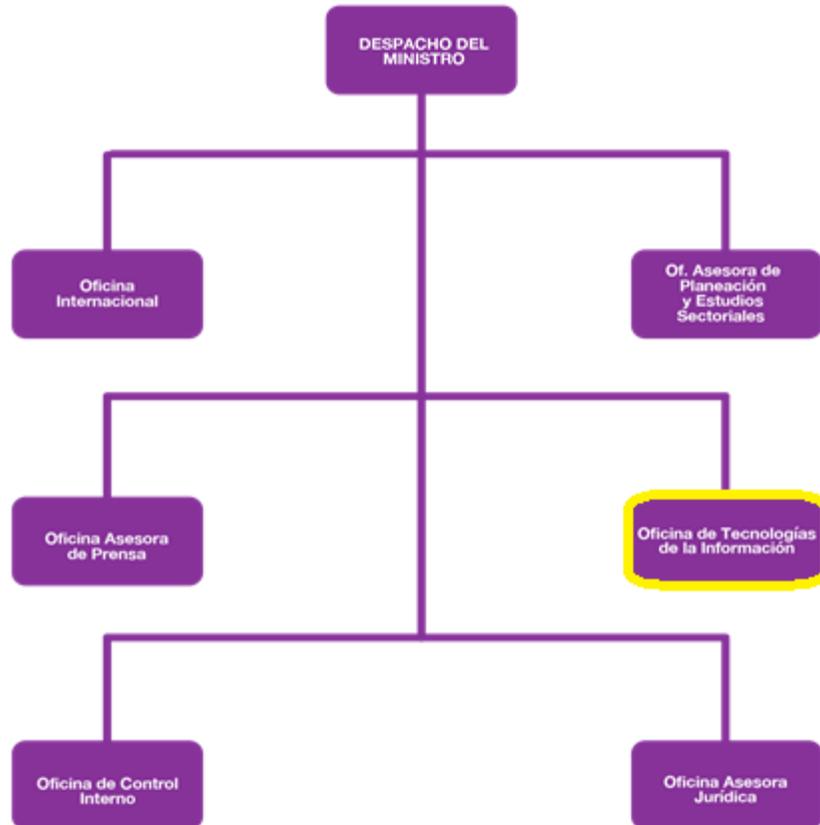


Figura 7 Estructura organizacional del Ministerio de Tecnologías de la Información y Comunicaciones, destacando la Oficina de Tecnologías de la Información

Fuente: ([MinTIC], 2019g)

La oficina de Tecnologías de la Información hace parte de los macroprocesos estratégicos de MinTIC, dependiendo directamente del despacho del Ministro(a) designado, tal y como se ilustra en el organigrama, a fin de garantizar que las tecnologías estén alineadas con la estrategia de la entidad y el sector.

Entre las funciones más relevantes de la oficina y que se encuentra en relación con el marco de este proyecto, encontramos que: se encarga de garantizar la aplicación de los estándares, buenas prácticas y principios para la información Estatal, definir lineamientos tecnológicos para el cumplimiento de estándares de

seguridad, privacidad, calidad y oportunidad de la información del Sector, vigilar que en los procesos tecnológicos de la entidad y del Sector se tengan en cuenta los estándares y lineamientos dictados por el Ministerio, que permitan la aplicación de las políticas que en materia de información expidan las entidades de acuerdo con su competencia; entre sus funciones también se encuentra apoyar al DANE en la definición de políticas, principios, lineamientos, e implementar los estándares y buenas prácticas que rigen la información estadística para lograr altos niveles de eficiencia en la producción de información estadística del Sector (MinTIC, 2019e).

Es así, como desde la oficina de tecnologías información, se definen las políticas, procedimientos y lineamientos entorno a la gestión de información en la entidad y el sector. El siguiente organigrama adaptado, ilustra las dos áreas que intervienen en este proyecto, la oficina de TI como proponente de la solución y la Dirección de Apropiación como área de ejemplificación.



Figura 7 Estructura organizacional adaptada del Ministerio de Tecnologías de la Información y Comunicaciones
Fuente: Elaboración propia

2.2 Marco normativo

A continuación, se presenta el marco normativo colombiano aplicable a la calidad de datos.

2.2.1 Normatividad Colombiana decreto 1008 de 2018 política de gobierno digital

La política de Gobierno Digital se estableció mediante el Decreto 1008 de 2018 (cuyas disposiciones se compilan en el Decreto 1078 de 2015, “Decreto Único Reglamentario del sector TIC”, específicamente en el capítulo 1, título 9, parte 2, libro 2), forma parte del Modelo Integrado de planeación y Gestión (MIPG) y se integra con las políticas de Gestión y Desempeño Institucional en la dimensión operativa de Gestión para el Resultado con Valores, que busca promover una adecuada gestión interna de las entidades y un buen relacionamiento con el ciudadano, a través de la participación y la prestación de servicios de calidad.

El anteriormente llamado Manual de Gobierno en Línea, ahora se denomina Manual para la Implementación de la Política de Gobierno Digital, entrega definiciones, lineamientos, estándares y acciones a ejecutar por parte de las entidades del Estado para dar cumplimiento con los objetivos propuestos en dicha política. Este manual fue creado por el Ministerio de Tecnologías de la Información como cabeza de sector en coordinación con el Departamento Nacional de Planeación (DNP) y se compone de cuatro pasos o momentos para su implementación. Ver figura.

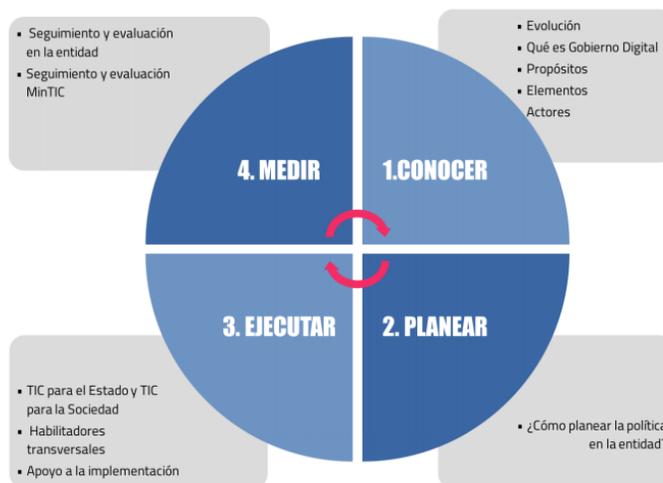


Figura 8 Momentos de la Política de Gobierno Digital
Fuente: (MinTIC, 2019, p. 7)

De esta manera, el gobierno mediante este decreto busca impulsar la implementación de estrategias digitales en las entidades del orden nacional y territorial en todo el país, para promover el uso y aprovechamiento de las

Tecnologías de Información y las Comunicaciones con el fin de brindar una comunicación más transparente entre el Estado y los ciudadanos (Gobierno Digital, s.f.).

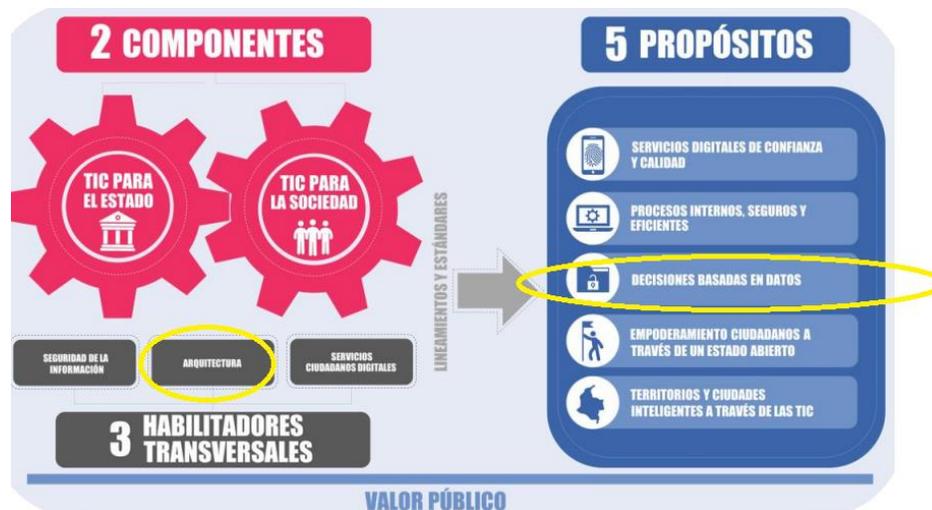


Figura 9 Diagrama conceptual de la política de Gobierno Digital.
Fuente: (MinTIC, 2019, p. 17)

En este diagrama, se pueden observar cada uno de los aspectos que intervienen en la política de Gobierno Digital, sus propósitos, componentes y habilitadores. Se resalta en esta imagen, el habilitador “**Arquitectura**” y el propósito “**Decisiones Basadas en Datos**” sobre el cual se basa el presente proyecto, Arquitectura para cumplir el propósito de toma de decisiones basada en datos.

2.2.2 Marco de referencia de arquitectura TI

Continuando en relación con lo planteado en el punto 2.2.1, en donde se referencia a la *Arquitectura* como un habilitador de la política y de acuerdo al Manual de Gobierno Digital, quien ejecuta es el grupo de trabajo de Arquitectura Empresarial, que actúa como comité técnico y evalúa los impactos de las decisiones de inversión en adquisición tecnológica, infraestructura, sistemas de información etc., esto depende de los niveles de madurez y necesidades de la organización en materia de arquitectura empresarial. De acuerdo con el manual de implementación de la política, este grupo de trabajo debe estar conformado por el director de Tecnologías de Información (CIO), el director de planeación, profesionales de arquitectura de sistemas de información, arquitecturas de infraestructura y el líder de Gestión de Información o arquitecto de información.

El habilitador de Arquitectura busca fortalecer las capacidades de gestión de T.I. de las entidades públicas, a través de la definición de lineamientos, estándares y mejores prácticas contenidos en el Marco de Referencia de Arquitectura Empresarial del Estado (Gobierno Digital, s.f.)

El Marco de la Arquitectura Empresarial es la estructura que ordena los conceptos y las estrategias, la columna vertebral del uso de tecnología, sobre la que las instituciones y los gobiernos soportan la gestión de TI. La Arquitectura TI permite que el Estado sea más eficiente al coordinar e integrar los esfuerzos de sus entidades y se basa en el Marco de Referencia que guía la gestión TI. (Arquitectura TI, s, f.). Este marco, se compone de seis dominios: Estrategia TI, Gobierno TI, Información, Sistemas de Información, Servicios Tecnológicos y Uso y Apropiación, ámbitos que agrupa los lineamientos, roles, normatividad, indicadores de medición.



Figura 10 Cuadro de los seis dominios del Marco de Referencia de AE.

Fuente: ([MinTIC], s.f.)

En la imagen se resalta el dominio sobre el cual está basado el presente proyecto, el dominio de Información.

El dominio de información es el responsable de diseñar los servicios de información, la gestión del ciclo de vida de los datos, el análisis de la información y el desarrollo de capacidades para el uso estratégico de la misma (MinTIC, 2018). Este dominio está conformado por 6 ámbitos:

- Planeación y Gobierno de los Componentes de Información.
- Diseño de los Componentes de Información.

- Análisis y aprovechamiento de los Componentes de Información.
- Calidad y Seguridad de los Componentes de Información.

2.2.3 Ley 1581 del 2012 Protección de datos personales

La ley 1581, complementa la regulación vigente para la protección del derecho fundamental que tienen todas las personas naturales a autorizar la información personal que es almacenada en bases de datos o archivos, así como su posterior actualización y rectificación.

Para su correcta implementación, todas las entidades de gobierno y empresas de sector privado, deben manejar con claridad los conceptos de autorización para recolectar y procesar los datos, la política de tratamiento de datos personales en la cual se contemplan los avisos de privacidad, medio de difusión, procedimientos para el tratamiento y medidas de seguridad, los derechos de los titulares de dichos datos y el manejo de transferencias y transmisiones internacionales de los datos (Ministerio de Industria Comercio y Turismo, 2013).

Dado el contexto de este proyecto, el cual se basa en análisis, diagnóstico y procedimientos para la gestión de la calidad de los datos, esta ley es aplicable dentro de la normativa que lo enmarca.

2.2.4 Ley 1712 del 2014 Transparencia

Esta ley busca promover y regular el derecho de acceso a la información de carácter público a los ciudadanos, mediante procedimientos y excepciones a la publicidad de la información, es decir todas las entidades del orden nacional y territorial, deben exponer al público la información que gestionan y que sea de interés para los ciudadanos, siempre y cuando ésta no esté considerada como reservada, para ello se basan en principios tales como: Transparencia, buena fe, facilitación, no discriminación, gratuidad, celeridad, eficacia, divulgación proactiva, responsabilidad en el uso de la información, y el principio que atañe a este trabajo, y por el cual se enmarca normativamente en esta ley, es el principio de la *calidad de la información* y que se define a continuación:

Ley **1712** "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" (Congreso de la República, 2014).

Principio de la calidad de la información: toda la información de interés público que sea producida, gestionada y difundida por el sujeto obligado, deberá ser oportuna, objetiva, veraz, completa, reutilizable, procesable y estar disponible en formatos accesibles para los solicitantes e interesados en ella. Para dar cumplimiento con este principio, es necesario que los datos procesados por las entidades y para el caso MinTIC, debe cumplir con las características de calidad ya mencionadas (MinTIC, 2014).

2.2.5 Directiva Presidencial de 2019

Esta directiva presidencial publicada el 02 de abril de 2019, tiene como objetivo principal promover el avance de la transformación digital del Estado impactando positivamente la calidad de vida de los ciudadanos a través de las interacciones digitales con el Estado. Esto quiere decir que, mediante la implementación del Portal Único del Estado Colombiano, se busca la integración de todos los trámites, servicios, información pública ejercicios de participación, colaboración y control social que las entidades públicas del orden nacional y territorial ofrecen a los ciudadanos (República de Colombia, 2019).

Se contempla el Estado como plataforma y para ello, el Ministerio TIC incorpora en el Manual de Gobierno Digital, los estándares y lineamientos que deben tener en cuenta las entidades públicas para integrarse al portal. Dentro de estos lineamientos, se destacan dos:

Integración de la información: se define la estandarización de información mínima a publicar bajo las definiciones de la presente directiva, la resolución 3564 de 2015 y la ley 1712 de 2014.

Integración a Datos Abiertos: se define como indispensable que las entidades que publican los datos dispongan de datos de calidad, pues de esto depende que el aprovechamiento de estos sea mayor (MinTIC, 2019).

2.2.6 CONPES 3920 de 2018 Explotación de Datos

El CONPES es la máxima autoridad nacional de planeación y se desempeña como organismo asesor del Gobierno en todos los aspectos relacionados con el desarrollo económico y social del país. Sus siglas traducen, Consejo Nacional de Política Económica y Social y fue creado por la ley 19 de 1958 (Departamento Nacional de Planeación [DNP], 2019a).

El CONPES 3920 busca el aumento del aprovechamiento de los datos en Colombia, a través del desarrollo de capacidades que les permita a las entidades del Estado gestionar los datos como activos de la organización, a fin de generar valor social y económico. Uno de sus objetivos es lograr la masificación de la disponibilidad de los datos para que sean accesibles, usables y de calidad, para ello, se plantea la línea de acción que diseña e implementa la infraestructura de datos de las entidades del Estado (Consejo Nacional de Política Económica y Social [CONPES], 2018a).

Otra línea de acción a resaltar de esta política es la de Institucionalizar la explotación de datos en la toma de decisiones públicas, mediante metodologías para la formulación de proyectos, la elaboración de un plan de Anticorrupción para incluir el aprovechamiento de datos en la lucha contra la corrupción y la emisión de guías que incentiven la explotación de datos (Consejo Nacional de Política Económica y Social [CONPES], 2018b).

2.2.7 Transformación Digital DNP

El Estado colombiano es consciente del proceso de transformación de la sociedad que atraviesa en la actualidad, y para ello, se plantea el CONPES de transformación digital que pretende promover la inversión y el desarrollo económico en el país a través de la innovación y el emprendimiento.

Esta política de transformación, consciente de las barreras que existen para su avance, busca atacar la falta de cultura, el desconocimiento en la adopción de tecnologías, potencializar negocios a base de tecnologías digitales, promover el uso y aprovechamiento de herramientas tecnológicas, fortalecer los marcos normativos para favorecer su adopción y proyectar estrategias para la 4RI y la industria 4.0. (DNP, 2019); precisamente y dentro de todos estos objetivos, se encuentran dos aspectos clave para su realidad, el uso de las tecnologías y los datos que cambian las actividades humanas, tanto así, que los expertos ya comienzan a plantear temas de calidad de datos no sólo en lo que se conoce como small data, sino que también

se empiezan a plantear escenarios de calidad de datos para la Big Data (Cai & Zhu, 2015e).



Figura 11 Diagrama de efectos económicos y sociales
Fuente: DNP (2019)

2.3 Marco conceptual y revisión de literatura

Los datos son un término general que denotan una representación simbólica (numérica, alfabética, algorítmica etc.) de algunos o todos los hechos que describen a un objeto, idea, situación, condición o cualquier otro factor. El dato por sí solo carece de sentido; sin embargo, al ser parte de un contexto y ser procesado de manera adecuada, se convierte en fundamental según (PowerData, s.f.a) para todos los niveles de la organización al punto de convertirse en uno de los activos más valiosos que soportan la toma de decisiones. A continuación, se relacionan los conceptos principales relacionados con calidad de datos.

2.3.1 Calidad de datos

La calidad de datos hace referencia a la conformidad que los datos presentan acorde con el propósito para el cual son designados, por lo que las organizaciones bajo este concepto analizan si estos son confiables, consistentes, si se encuentran actualizados y si están libre de duplicidades (Fisher, 2009).

Si bien el concepto por su definición pudiera parecer ambiguo, la calidad de datos como dice Loshin (2001) puede ser cuantificable, medible y mejorada manteniendo siempre el enfoque hacia un retorno de inversión. Esto quiere decir, que la gestión que se realice sobre los datos debe estar alineada siempre hacia los objetivos estratégicos de las organizaciones para que sean altamente

aprovechables por la alta gerencia. Contar con datos de alta calidad permite generar información confiable para toma de decisiones de tal modo que, en el contexto de la Dirección de Apropiación, podrían conocer y determinar cuáles son las zonas del país mayormente beneficiadas con sus programas, que población se está impactando y determinar cuáles programas ofrecer o no en una región determinada.

2.3.2 Gestión de calidad de datos

La gestión de calidad de datos consiste en aplicar los conceptos y prácticas de la Gestión de la Calidad Total (GTC) para mejorar la calidad de los datos y la información, lo que implica que se deban definir políticas y normas, realizar el análisis del estado actual de los datos, realizar limpiezas y la corrección de los datos, la mejora de la calidad de los datos y la educación hacia los diferentes actores (Santos & Lucas, 2019).

En el marco de la gestión de calidad de datos, las dimensiones de calidad son necesarias para medir la calidad de los datos y poder realizar acciones sobre los hallazgos que se encuentren de su evaluación; para ello el DAMA (2017) presenta las dimensiones “Core” que se listan a continuación para la identificación de los niveles de calidad de los datos:

- **Exactitud:** la exactitud de datos se refiere a la medida en que los datos representan correctamente las entidades de la “la vida real” que estos modelan. Los datos de la Base pueden ser comparados con otros de una base de datos similares y deben ser los mismos.
- **Compleitud:** deben existir valores asignados en el dato. Todas las filas del conjunto de datos deben estar presentes. Asignación de atributos de restricción u obligatoriedad.
- **Consistencia:** los valores en un conjunto de datos son consistentes cuando no exceden el rango de atributos previamente parametrizados, por ejemplo, se podría esperar que el número de transacciones por día no exceda el 105% del promedio de los 30 días anteriores.
- **Integridad:** se refiere a que todas las referencias o relaciones de la tabla a la que pertenece el dato (campo) son válidas en otros conjuntos de datos con los que se pueda relacionar.

- **Unicidad:** ninguna instancia de entidad se registrará más de una vez con las mismas características. Hace referencia a que ninguna entidad de datos exista más de una vez dentro de un conjunto de datos.
- **Validez:** el dato es válido si está conforme al tipo de dato esperado, rango o estándar según se haya definido. Se refiere a si las instancias de datos se almacenan, intercambian, o si se presentan en un formato que es válido para el dominio de los valores.
- **Oportunidad:** se refiere a la expectativa de accesibilidad y disponibilidad de la información. Están disponibles cuando se requiere.

2.3.3 Gobierno de datos

Según (Thomas, s.f.) en un artículo del Data Governance Institute, el gobierno de datos se define como un sistema de derechos y responsabilidades de decisión para procesos relacionados con la información y se ejecutan con base en los modelos diseñados previamente en los que se describe quienes pueden tomar determinadas medidas, sobre qué información, cuándo, y bajo qué métodos y/o herramientas.

En otras palabras, es una estrategia que le permite a las organizaciones generar valor a partir de los datos concibiéndolos como un activo fundamental, esta estrategia permite ejercer autoridad y control (planeación, monitoreo y cumplimiento) sobre la gestión de datos en el negocio y para su ejecución requiere de 3 componentes clave como son: las personas, los procesos y la tecnología.

La estrategia de gobierno de datos está encaminada a planificar el modelo de gobierno de datos que el negocio requiere para tratar los datos como un activo de negocio. Para definir y establecer la estrategia de gobierno de datos es necesario establecer políticas, procesos, roles, responsables y tecnología basado en las mejores prácticas para gestión de datos.

El concepto de gobierno de datos es suficientemente amplio, sin embargo, modelos de referencia de gestión de datos como el DAMA definen que un modelo de gobierno de datos que respalde la estrategia de negocio mínimo debe incluir:

- Alineación estratégica: impulsores, misión y visión.
- Objetivos, principios y políticas.
- Procesos relacionados con la operación de los datos.
- Reglas de negocio.

- Roles y responsabilidades.

Como parte de los objetivos del gobierno de datos, se destacan, el asegurar que los datos cumplen con las demandas del negocio, reducir los costos en la gestión de los datos, proteger los datos que se consideren sensibles, gestionar y desarrollar los datos como un activo estratégico y de valor (IBM, 2012).

Para la revisión de literatura, se consultaron diferentes fuentes bibliográficas a fin de conocer lo que dicen los expertos del tema desde su experiencia en las organizaciones y de esta manera revisar los fundamentos teóricos para proponer las buenas prácticas en gestión de calidad de datos requeridas en la Dirección de apropiación de MinTIC.

2.3.4 Referencias de literatura sobre calidad de datos

La calidad de un dato se puede definir según Thomas C. Redman (2001) destacado experto en calidad de datos, de la siguiente manera: "Los datos son de alta calidad si son aptos para sus usos previstos en operaciones, toma de decisiones, análisis y planificación." (Discover The New, 2017, p.1)

Wade, Loucks, Macaulay, & Noronha (2018a) plantean que una de las capacidades de la agilidad digital es la toma de decisiones informada. Los datos se analizan y distribuyen para soportar decisiones estratégicas y reglas de negocio automatizadas que dependen de la cantidad y calidad de los datos recogidos, también destaca que Velthuis (2018c) "datos malos hacen imposible una buena toma de decisiones" (p. 26).

En la década de 1950, los investigadores comenzaron a estudiar problemas de calidad, especialmente para calidad de productos. Puntualmente la investigación sobre calidad de datos se da en la década de los 90', época en la que diferentes académicos propusieron varias definiciones de calidad de datos y empezaron a plantear temas como las dimensiones de los datos. Así, por ejemplo, el grupo de Gestión de Calidad de Datos del MIT en su investigación sobre calidad de datos, la define como la "aptitud o idoneidad para el uso" (Cai & Zhu, 2015d) y proponen que esta definición está sujeta a las necesidades de los consumidores de los datos, a su vez también definen que las dimensiones corresponden a un conjunto de características que la representan.

Por su parte, Pipino, Lee, Wang, & Yang (2002) afirman que los estudios realizados, demuestran que la calidad de los datos es un concepto multidimensional y que por lo tanto, las empresas deben tratar tanto con las percepciones subjetivas

de los individuos involucrados, como con las mediciones objetivas que se basan en preguntas establecidas que responden a la calidad de datos. Dichas mediciones parten del entendimiento de los requisitos que deben cumplir los datos para ser de calidad, este conjunto de características de los datos por ejemplo, (integridad, validez, exactitud, consistencia, disponibilidad y puntualidad) definen que los datos cumplen o no con los requisitos (International Organization for Standardization [ISO], s.f.).

En China, hacia el año 2008, se creó un grupo de investigación sobre calidad de datos, y empezaron a discutir temas como la definición de lo que es calidad de datos, sus fuentes de error y las posibles acciones de mejora. Los enfoques cuantitativos y cualitativos integrados a los problemas de calidad a través de la aplicación del principio de Pareto indican que, el 80% de los problemas que se presentan, provienen de un 20% de las causas (Keller et al., 2017). Otra perspectiva que plantea Keller consiste en tratar de eliminar los errores para mejorar la eficiencia y la calidad de los resultados extendiendo las definiciones a los ámbitos operativos, esto quiere decir, que la calidad de los datos se podría clasificar en los siguientes tres tipos:

- Calidad operativa: que examina el trabajo estratégico y táctico.
- Calidad del comportamiento: que se centra en los aspectos humanos y las actividades cotidianas.
- Calidad del proceso: que analiza técnicas y métodos que incluyen herramientas de control de datos y sistemas de información.

Por último, es importante resaltar lo que menciona Keller acerca de la clasificación de los tipos de datos que manejan las aplicaciones de hoy, los cuales son: datos diseñados, es decir datos que han sido recolectados mediante encuestas, experimentos y otros instrumentos de investigación. Datos de oportunidad, así los denomina para referirse a los que son producidos a través del avance tecnológico por el uso de la tecnología en la cotidianidad de las personas como los sistemas de GPS, redes sociales, dispositivos móviles y accesos a internet entre otros. Y los datos administrativos que corresponden a las actividades específicas sobre las personas o entidades Keller et al. (2017), tales como (fechas de nacimiento, matrimonios, actividad industrial, bancaria entre otros), en otras palabras, son todos los registros resultantes de las necesidades fiscales, tributarias u otras actividades para la administración y seguimiento del funcionamiento de las entidades públicas o privadas y de los programas que implementan los gobiernos (Departamento Nacional de Estadística [DANE], s.f.).

Durante la segunda reunión de la Conferencia Estadística de las Américas de la CEPAL, en Santiago de Chile en 2003, se llevaron a cabo discusiones acerca de la calidad de datos, los registros administrativos y la credibilidad. En dicha

conferencia, se resalta la importancia que pueden llegar a tener los datos administrativos si cuentan con la calidad suficiente para ser usados como fuentes de información primarias, en lugar del uso de herramientas como censos y las encuestas, “En los países estadísticamente más avanzados se ha ampliado el concepto tradicional de calidad y dentro de él se incluyen atributos o dominios generalmente aceptados, referidos a: relevancia, exactitud, oportunidad y puntualidad, accesibilidad y claridad, comparabilidad, coherencia y completitud.” (Echegoyen, 2003, p.15). Esta definición de dominios o características, coinciden con las planteadas por otros autores que se han referido a este mismo tema. Es así como Suecia es uno de los países europeos que ha logrado sacar provecho de estos datos administrativos con calidad, para fines estadísticos (Wallgren & Wallgren, 2016).

Todo lo anterior permite concluir que los datos que poseen las empresas son de vital importancia para las organizaciones, tal como lo mencionan Westerman et al. (2014b) en su libro *Leading Digital* “los datos son el tipo de activo prominente en el mundo digital y requieren una atención muy significativa por parte de los ejecutivos. Los datos se han convertido en los activos digitales más valiosos para las empresas” (p. 182).

Con base en la anterior revisión, en la siguiente sección se va a presentar en detalle el marco de referencia DAMA y de otros propuestos por grandes consultores en el tema.

2.3.5 Marcos de referencia para la gestión de calidad de datos.

A continuación, se describe lo que dice el DAMA acerca de los puntos más importantes de la calidad de datos encontrados en el proceso de revisión de literatura, también se encontró literatura desarrollada por importantes consultoras para la gestión de proyectos de calidad de datos producto de la experiencia en el tema, los cuales serán tenidos en cuenta durante el desarrollo de este trabajo.

2.3.6 DAMA

El DAMA por sus siglas en inglés [Data Management Association] es una organización de reconocimiento internacional, que se encarga de la definición y formalización de estándares y buenas prácticas para la gestión de datos, así como de promover su entendimiento y desarrollo. Para ello, han desarrollado un documento titulado: Guía para la Entidad de Conocimiento de la Gestión de Datos

[Guía DAMA-DMBOK], la primera versión salió en Julio de 2016 y la versión más reciente fue publicada en el año 2017 (DAMA, 2017), sobre la cual se desarrolla el presente trabajo.

En el siguiente diagrama, se puede observar cómo se encuentra estructurado el marco de referencia que propone el DAMA para todo el proceso de gestión de información, en el cual se resalta el aspecto objeto de esta investigación; la calidad de datos.



Figura 12 Marco para la Gestión de Datos del DAMA
Fuente: Elaboración Propia

La estrategia de este marco de referencia se basa en diez elementos de gestión: la arquitectura, modelamiento, almacenamiento, seguridad, integración, gestión de contenidos, datos de referencia o también conocidos como datos maestros, inteligencia de negocios, metadatos y calidad de datos, si bien son importantes todos estos aspectos para gestionar el ciclo de vida del dato, nos concentramos en la calidad de datos.

Marilú López, presidenta del DAMA en México manifiesta que, “Sabemos la importancia de tecnologías de punta, pero también estamos conscientes que atrás de las tecnologías necesitamos forzosamente tener una disciplina formal para el manejo, para la interpretación y la mejor explotación de los datos, que nos permita realmente tratarlos como activos” (InfoChannel, 2017).

Volviendo al tema específico, dentro del elemento de calidad de datos, el DAMA plantea las actividades a desarrollar, herramientas, técnicas, gobernanza y guías para la implementación, iniciando por un diagnóstico, siempre teniendo en cuenta aspectos ambientales de la organización. A continuación, se presentan los puntos más relevantes de este elemento:

Elementos Calidad de Datos - DAMA 	
	Helpful
1. Actividades	 <ul style="list-style-type: none"> Definir datos de alta calidad  Definir una estrategia de calidad de datos Identificar datos críticos y reglas de negocio Realizar una evaluación inicial de la calidad de datos Identificar y priorizar las posibles mejoras Definir objetivos para la mejora de la calidad de los datos Desarrollar e implementar operaciones de calidad de datos
2. Herramientas	 <ul style="list-style-type: none"> Herramientas de perfilado de datos Herramientas de consulta de datos Herramientas de modelado y ETL Plantilla de reglas de calidad de datos Repositorios de Metadatos
3. Técnicas	 <ul style="list-style-type: none"> Acciones preventivas Acciones correctivas Módulos de control de calidad y código de auditoría Métricas efectivas de calidad de datos Control estadístico de procesos Análisis de causa raíz
4. Guía de Implementación	 <ul style="list-style-type: none"> Evaluación de la preparación / Evaluación de riesgos Organización y cambio cultural
5. Calidad y Gobierno de los datos	 <ul style="list-style-type: none"> Política de Calidad de Datos Métricas

Tabla 1 Elementos de calidad de datos.
Fuente: Elaboración propia.

Antes de iniciar con cualquier programa de calidad de datos en las organizaciones, es importante que todos los involucrados conozcan y entiendan el significado de calidad del dato, conocer las necesidades del negocio, identificar los puntos de dolor de la organización y empezar a construir en consenso con todos, las prioridades y controles para el mejoramiento de la calidad de los datos (DAMA, 2017b). Lo anterior, se encuentra en el marco de la sección de actividades propuestas por esta guía, así como las demás relacionadas en la tabla 1, tales como la definición de una estrategia que debe estar alineada con las estrategias del negocio, la identificación de datos críticos y reglas de negocio, un diagnóstico de calidad de datos, identificación y priorización de mejoras, definición de objetivos

para las mejoras a realizar y la puesta en marcha o desarrollo en los procesos de información.

En la sección de herramientas se plantea que, si bien para que las organizaciones inicien muchas de estas, en ocasiones las herramientas ya contienen de forma predefinida paquetes de reglas para ser implementadas. Sin embargo, las empresas necesitan definir y crear sus propias reglas de acuerdo con el contexto y necesidades particulares de cada una de ellas (DAMA, 2017c).

Para ello se sugiere usar acorde al contexto, el perfilamiento de datos, el cual permite la identificación de patrones y entrega una evaluación inicial de las características de los datos, seguida de esta se pueden implementar los querying data o consultas a las bases de datos, que ayudará a la identificación y cuantificación de problemas más específicos tales como registros duplicados e integridad de los datos. Estos pueden ser producto de lo que se logró identificar en los patrones de datos encontrados en el perfilamiento.

Dentro de las herramientas para el mejoramiento de la calidad encontramos las ETL [Extracción, Transformación, Carga] y las plantillas de reglas de negocios. Estas dos se deben usar en conjunto con las áreas de negocio y el equipo técnico de desarrollo, de tal forma que las necesidades sean cubiertas desde los dos aspectos. A estos instrumentos se debe unir la definición de metadatos para asegurar la calidad de los requerimientos, reglas, mediciones, y documentación, para que sean válidas y entendidas por los usuarios que consumen los datos (DAMA 2017d). Por otra parte, las acciones preventivas que son menos costosas de realizar, acciones correctivas que elevan el costo de operación, módulos de control de calidad, mediciones, control estadístico de procesos y análisis causa-raíz, hacen parte de la sección de técnicas que recomienda el DAMA para realizar en un programa de calidad de datos.

Mejorar la calidad de los datos de una organización por pequeña que ésta parezca, no es una labor sencilla de realizar, aun cuando éste sea lanzado como un programa de gobierno de datos y se cuente con profesionales capacitados para ello.

Mejorar la calidad de los datos requiere cambios acerca de cómo las personas deben pensar y comportarse frente a los datos, en otras palabras, requiere de un cambio cultural que demanda planeación, entrenamiento y refuerzo que requieren ser medidos durante el paso del tiempo, estos aspectos se encuentran en la sección de guía de implementación (DAMA 2017e).

Por último, pero no menos importante, se encuentra la definición de políticas y métricas de medición para mantener una buena calidad de datos en la organización. En ocasiones, los programas de calidad de datos que se desarrollan en las empresas dan paso a la creación y establecimiento de áreas o programas de

gobierno de datos, con los que no se contaba, a fin de garantizar que el esfuerzo que se haya realizado sea controlado desde dicha área a través de los procesos de las compañías. Las buenas prácticas para esto están detalladas en la sección de calidad y gobierno de datos (DAMA 2017f).

2.3.7 ISO 8000

ISO [International Organization for Standardization], es un organismo mundial de estandarización que desarrolla normas internacionales voluntarias. Nace en el año 1946 y ha publicado más de 20.000 normas de diferentes sectores de ámbitos tecnológicos y de negocios (International Dynamic Advisors [Intedya], s.f.). La alta competencia internacional acentuada por los procesos globalizadores de la economía y el mercado, junto con la importancia que toma la opinión de los consumidores cada vez más, ha provocado que estas normas pese a su carácter voluntario hayan ganado reconocimiento y aceptación internacional (ISOTools, 2015a).

Dentro de la finalidad de las normas ISO, se encuentra el ofrecer orientación, coordinación, simplificación y unificación de criterios a las empresas y organizaciones para lograr la reducción de costos y aumentar la productividad. Entre algunas de sus ventajas están:

- Proporcionar elementos para alcanzar y mantener mayores niveles de calidad en los productos y/o servicios.
- Ayudar a la satisfacción de los clientes.
- La implementación de estas normas se constituye en una ventaja competitiva.
- Mejor y mayor acceso a mercados internacionales (ISOTools, 2015b).

Entre algunas de las familias de las normas ISO, se encuentra la ISO 8000; esta hace referencia a la calidad de datos y está compuesta por 4 partes principales.

- Conceptos generales y medición.
- Procesos de Gestión de calidad de datos.
- Intercambio de datos maestros.
- Información de ingeniería.

En la siguiente figura se resalta lo que propone la ISO relacionado con la gestión de calidad de datos.

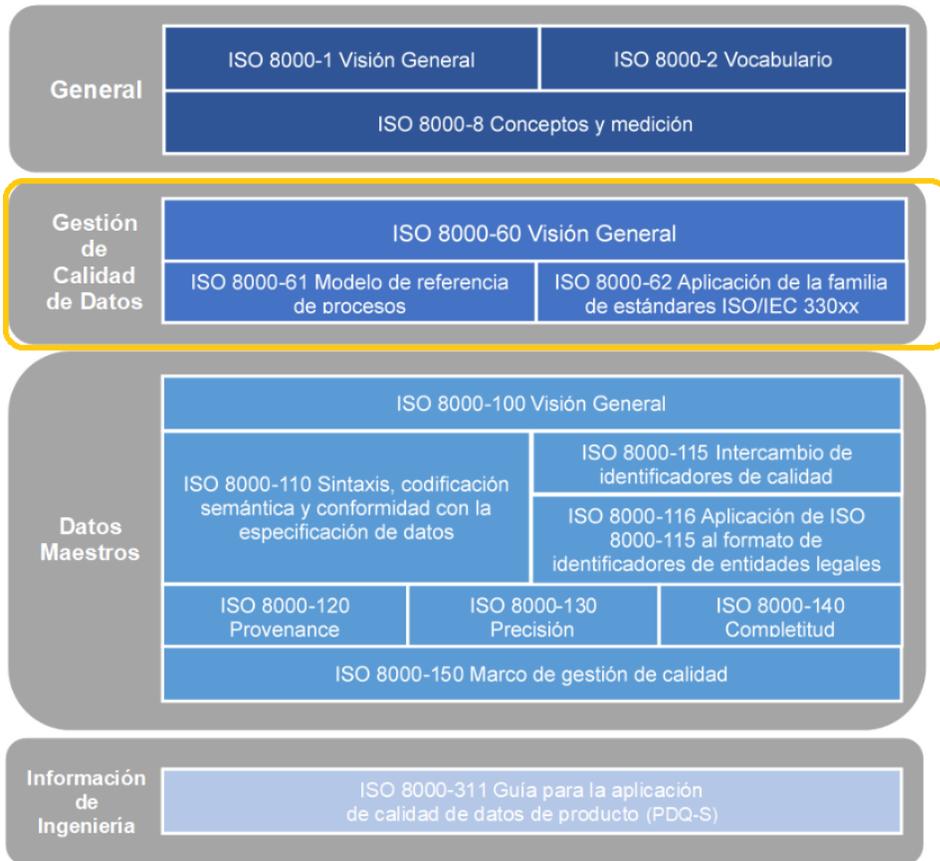


Figura 13 Mapa general de la norma ISO 8000
Fuente: (ISO 8000, 2018)

La parte que se enfoca en la calidad de datos, según este estándar, está en la parte 100, teniendo como impulsores principales, el mejoramiento de la calidad de dato, permitir la integración para construir términos de relacionamiento con otras entidades.

La parte 110, describe el uso de la sintaxis, codificación semántica y presenta los requerimientos. La sintaxis se refiere a las reglas formales para gobernar la construcción de sentencias válidas en un lenguaje determinado. Por otra parte la codificación semántica se refiere a la definición que le dan sentido a las sentencias de los lenguajes (Jager, 2016a).

A continuación, se relaciona la figura en la que se muestran la familia de estándares de la ISO 8000.



Figura 14 Familia de estándares de la ISO 8000
Fuente: (Jager, 2016b)

2.3.8 MIKE 2.0

MIKE 2.0 [Method for an Integrated Knowledge Environment] es una metodología de código abierto para la Gestión de Información Empresarial desarrollado por BearingPoint Consulting, que proporciona un marco para el desarrollo de la información que se puede aplicar a diferentes tipos de proyectos de Gestión de Información. El equipo de BearingPoint comenzó la iniciativa MIKE2 a principios de 2005.

Algunos de sus objetivos son: dirigir un enfoque general a través de la estrategia de una organización, habilitar las capacidades de las personas mientras se apropia una cultura de excelencia de información, mejorar los procesos de cumplimiento en torno a la información, políticas, prácticas y medición (Mundo BI, 2011).

En la siguiente figura, se ilustra el marco de la metodología que propone Mike y se resalta el módulo de calidad de datos que se encuentra dentro del proceso de Gestión de Datos Empresariales.

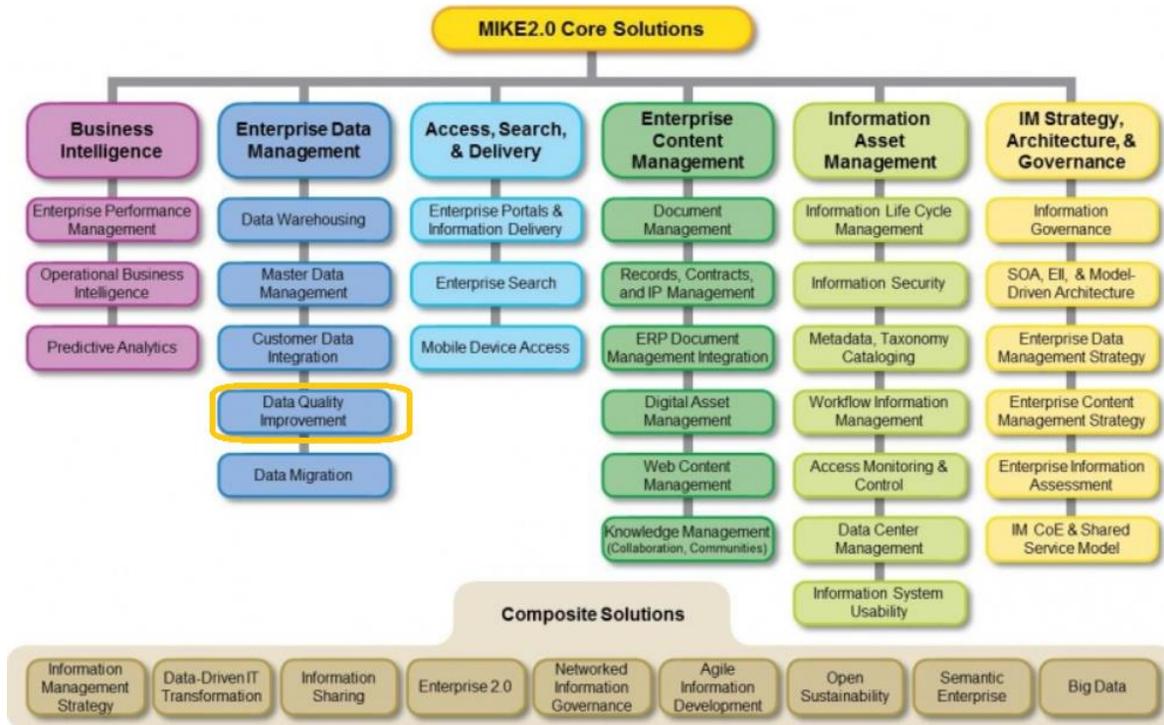


Figura 15 Marco de la metodología Mike 2.0
Fuente: (MIKE 2.0, s.f.-b)

La metodología contempla 5 fases para su implementación:

Fase 1: Evaluación del negocio. Constituye la base de la estrategia para toda la implementación.

Fase 2: Evaluación de la tecnología. Consiste en entender en dónde se encuentra la organización respecto a la tecnología y hacia dónde quiere llegar, con esto logra la identificación de brechas.

Fase 3: Gestión de la información, hoja de ruta.

Fase 4: Incremento del diseño.

Fase 5: Desarrollo Incremental.

Las últimas tres fases, de la 3 a la 5, se desarrollan de forma iterativa y de manera cíclica en marcos ágiles, dado a que la dinámica de esta metodología pretende entregar soluciones pequeñas con poca funcionalidad pero que, a medida que crecen las iteraciones, se van refinando hasta lograr el objetivo final del proyecto (AIIM International, 2009). Por cada una de estas fases, se encuentran una serie de actividades que se van desarrollando y describen las principales unidades de trabajo a través de un enfoque en conceptos clave y entregables.

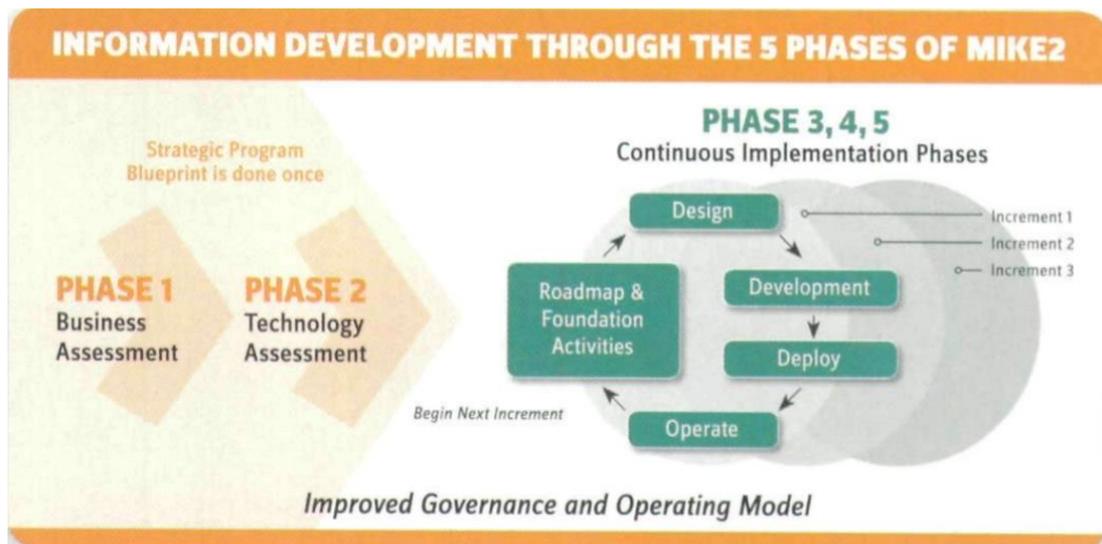


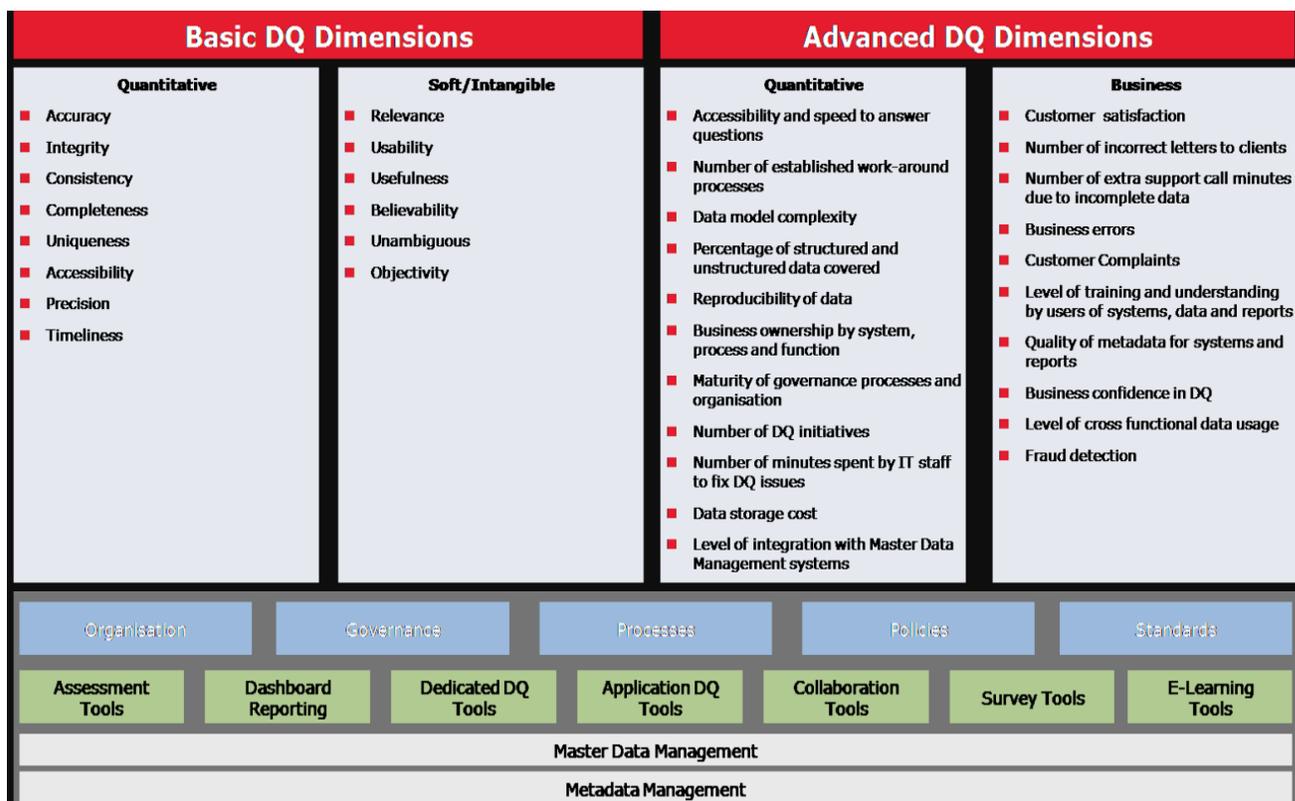
Figura 16 Fases de Implementación de la metodología Mike 2.0
Fuente: (AIIM International, 2009)

En la fase 3, hoja de ruta de gestión de la información, la metodología cuenta con una lista de actividades clasificadas, de las que se resaltan para el presente trabajo las siguientes:

Métricas de gobernabilidad de datos: en esta sección se proponen los objetivos en cuanto a la calidad de la información que la organización pretende lograr y se definen métricas a las que se les debe realizar seguimiento por parte de la alta dirección y de los conocedores de los datos, que han sido seleccionados como de alto impacto. Para ello, se deben realizar tres actividades:

1. Definir categorías para las mediciones a realizar sobre los conjuntos de datos.
2. Por cada conjunto de datos clave, recopilar un estado actual de la medición.
3. Definir las nuevas métricas que van a ser medidas sobre los conjuntos de datos, estas métricas deben estar alienadas con los objetivos estratégicos de la organización (MIKE 2.0, s.f.).

En la figura a continuación, se ven listadas las dimensiones de calidad de datos y la clasificación de las métricas que recomienda Mike 2.0 para evaluar los conjuntos de datos clave.



¿Cómo se miden las métricas de gobernabilidad de datos?		
Categoría métrica	Descripción	¿Cómo se mide la métrica?
Exactitud	¿Los datos representan con precisión la realidad o una fuente verificable?	Auditoría
Integridad	¿Existen enlaces rotos entre los datos que deberían estar relacionados?	Perfil / Reglas de Negocio
Consistencia	¿Hay una sola representación de datos?	Perfil / Reglas de Negocio
Lo completo	¿Falta alguna información clave?	Perfil / Reglas de Negocio
Validez	¿Se almacenan los datos en un formato aceptable y contienen valores válidos?	Perfil / Reglas de Negocio
Accesibilidad	¿Los datos son fácilmente accesibles, comprensibles y se utilizan de manera consistente?	Encuesta
Oportunidad	¿Se registra la información y se pone a disposición de los sistemas tan rápido como se requiere?	Encuesta

Figura 17 Enfoque para medir la Calidad de Datos y ejemplo de medición.
Fuente (MIKE 2.0, s.f.b)

Perfil de los Datos: consiste en realizar un análisis de la información basado en la evaluación a conjuntos y estructuras de datos reales, para identificar problemas de calidad a nivel de campos, tablas y entre tablas, capturar los metadatos que serán usados en el proceso de cargue de datos e identificar las reglas de negocio para el manejo de los datos. Esta sección se compone de las siguientes actividades:

1. Prepararse para la evaluación, es decir, contar con los ambientes y herramientas para la ejecución de la prueba y tener definido el alcance de los datos que se evaluarán.
2. Realizar perfilado de columnas, por cada entidad de datos, se deben evaluar los datos almacenados en cada columna.
3. Realizar perfiles de tablas, implica analizar los datos a nivel de filas de una sola entidad de datos.
4. Realizar perfilado de múltiples tablas para buscar problemas de redundancia e integridad referencial.
5. Realizar informe de calidad de datos, este se compone de los resultados de la evaluación realizada en cada paso y del planteamiento de recomendaciones acerca de si los datos deben continuar cargándose o no al sistema destino (MIKE 2.0, s.f.c)

Esta metodología presenta una ventaja frente a las tradicionales, y es que básicamente aprovecha su naturaleza colaborativa. Lo que significa que con el tiempo, el refinamiento y el aporte de las organizaciones, puede convertir a MIKE2.0 en una metodología sólida y consistente que proporciona buenas prácticas que son aceptadas dentro de la industria.

2.4 Buenas prácticas en proceso de calidad de datos

Dentro de las buenas prácticas en procesos de calidad de datos, se destaca el perfilamiento de datos y su evaluación que se describe a continuación.

2.4.1 Perfilamiento y evaluación de datos

El perfilamiento de datos es un proceso que hace parte de la calidad de datos y el almacenamiento de estos. Dicho perfilamiento de datos hace parte de un proceso previo o inmerso a la construcción de ETL (Extraction, Transformation and Loading). Viéndolo desde esta perspectiva, el perfilamiento detecta en los datos pequeñas anomalías que deberán ser eliminadas o mitigadas antes de iniciar el análisis de la información, y desde luego mucho antes de realizar el respectivo despliegue o publicación de los datos en los ambientes de producción (Kimball, 2004a).

¿En qué consiste el Perfilamiento?

El perfilamiento de datos corresponde a un análisis sistemático del contenido de una fuente de datos, esto se hace a partir de la comprobación y verificación de cada uno de los campos de una tabla de base de datos, con el objetivo de observar si este concuerda con los tipos de datos, formatos, definiciones básicas de los datos y de dominio.

Es de gran valor el perfilamiento para identificar:

- Cuántas filas tienen valores nulos, o tienen contenidos que violan la definición del dominio. Ejemplo, si la definición de dominio es “Año” quiere decir que el registro de valores alfanuméricos representa claramente un problema.
- Relaciones entre campos de la misma tabla con el fin de identificar cuál de ellos podría llegar hacer una llave primaria; esto es útil ya que la violación de registros únicos son errores que rompen las reglas de negocio.

Es importante aclarar que el perfilamiento solamente diagnostica y detecta anomalías y no corrige los datos (Kimball, 2004b).

Durante la evaluación de la calidad de datos se sugiere seguir los pasos descritos a continuación:

- Análisis de elementos (propiedades de columnas): se evalúan los valores digitados en cada columna con el objetivo de identificar si lo registrado es válido o no. Se evalúan tipos de datos, formatos, rangos etc.
- Análisis de la estructura (dependencias funcionales, reglas de integridad): en este paso se trata de identificar las dependencias funcionales en cada tabla, detectar pares de columnas que son sinónimos y representan el mismo valor y verificar reglas de integridad evaluando las llaves primarias y foráneas.
- Verificación de reglas de negocio (simples, compuestas): se aplican reglas de negocio al conjunto de datos con el fin de detectar las anomalías; si existen anomalías puede ser que hacen falta datos o que los que existen poseen errores que impiden cumplir con dicha regla de negocio.
- Análisis estadístico: Se logra complementar los análisis anteriores en los cuales no se pueden aplicar reglas de negocio. Realizando o generando estadísticas de los datos se puede obtener conteos, distribuciones, frecuencias, promedios, valores máximos y mínimos que ayudan a evaluar si los valores son lógicos o no.

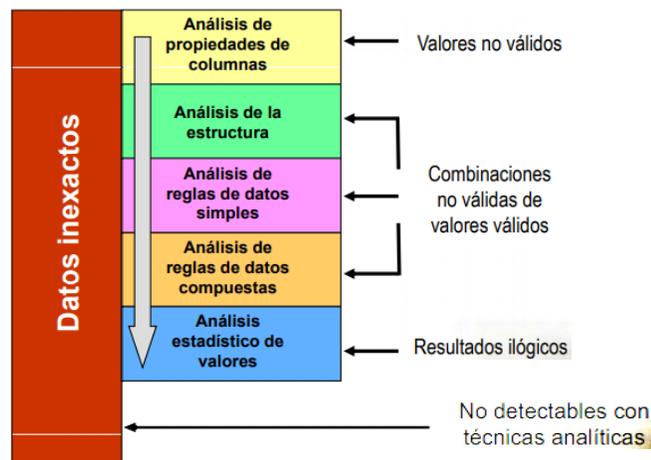


Figura 18 Diagrama de pasos para la evaluación de calidad de datos
Fuente: (Alvarado, 2011)

Una vez realizada la evaluación de la calidad de datos es importante, según (fluxicon, s.f.) empresa que ha estado al frente del movimiento del proceso de minería desde 2009, tener en cuenta los ítems descritos a continuación, los cuales se deben ir chequeando durante el proceso de calidad de datos con el fin de detectar si se ha realizado un buen proceso de calidad:

- No existen errores durante la importación de datos.
- No existen huecos en la línea de tiempo.
- Cantidad de datos esperada.
- La distribución de los valores se encuentra acorde con la realidad (no existen valores atípicos o inesperados).
- Documentación de todos los problemas presentados y preguntas sobre los datos.
- Si ha tenido que excluir datos debido a problemas de calidad de datos, ¿sigue siendo representativo el conjunto restante?

El conocer y comprender a un mayor nivel de detalle los datos, ayuda en la detección temprana de problemas, que harán que el desarrollo del proyecto se realice con mejores tiempos y cumpla a cabalidad los objetivos propuestos.

La conformación de un equipo de trabajo que se encargue de ejecutar las tareas de perfilamiento de datos hace que se cuente con personal que conoce lo que se está haciendo, conoce las reglas del negocio y que efectivamente ayuda a anticipar los complejos problemas de calidad de datos que se puedan presentar durante el proyecto. Como expresa Kimball (2004c) “El equipo evita las vergonzosas sorpresas de descubrir problemas cerca del cierre o al final de un proyecto” (p.2) .

3 Metodología

En este capítulo se presenta la metodología implementada para el desarrollo del presente trabajo de grado.

3.1 Diseño de la investigación

El tipo de enfoque metodológico utilizado en este proyecto es cualitativo, basado en que la finalidad de este enfoque como lo dice (Hernández Sampieri, Collado, & Baptista Lucio, 2010a) es: “Describir, comprender e interpretar los fenómenos, a través de las percepciones y significados producidos por las experiencias de los participantes” (p.11).

Teniendo en cuenta que los objetivos están orientados hacia el entendimiento y descripción de la problemática actual, a partir de sus involucrados directos y que como razón circunstancial todos los componentes de la metodología desarrollada como muestras e instrumentos de recolección de información, se basaron de acuerdo al conocimiento previo directo del investigador del fenómeno de estudio por su rol actual desde la Oficina de TI, el presente trabajo se desarrolla bajo una tipología de estudio cualitativa que permite el alcance de los objetivos propuestos; para ello, la propuesta metodológica de investigación se plantea en dos fases:

En la primera fase, se realiza un diagnóstico, donde se escoge bajo muestreo teórico una de las bases de datos de la Dirección de Apropiación y mediante la aplicación de una matriz elaborada por el autor del presente trabajo basada en las dimensiones señaladas en los marcos de referencia, se pretende soportar desde el aspecto técnico la justificación del problema.

Para la segunda fase, se define la aplicación de entrevistas a profundidad como instrumento de recolección de información, en donde las fuentes primarias se dividieron en dos grupos, cada uno con su guía de entrevista específica. El primer grupo, son los responsables del manejo de los datos en la Dirección de Apropiación, con los cuales se busca describir, comprender e interpretar la situación actual que se vive en torno a la problemática. El segundo grupo está compuesto por profesionales expertos en el tema de calidad de datos, con los que previamente en el rol que ejerce el autor desde la Oficina de TI han tenido alguna interacción en temas de gestión de datos dentro de la Entidad; de estos se enlistan seis (6) en total, de los cuales cuatro (4) fueron efectivos en la aplicación del instrumento de recolección de información seleccionado. Con ellos se pretende encontrar recomendaciones y buenas prácticas que permitan aportar a la construcción de la

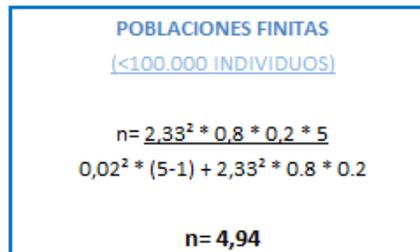
propuesta, mediante la triangulación de los marcos de referencia descritos en el presente trabajo y las respuestas a sus entrevistas.

Finalmente, de acuerdo con los resultados de las dos fases aplicadas, se pueden identificar los aspectos más relevantes para tener en cuenta en el planteamiento de la propuesta, que sirva como herramienta para garantizar un mejor desempeño en las labores diarias de la gestión de datos en la Dirección de Apropiación con base en el alcance definido.

3.1.1 Selección de la muestra

Acorde con lo descrito en el diseño de investigación, la selección de la muestra está dada de la siguiente manera:

Muestra 1: el tamaño de la muestra para las bases de datos determinado bajo fórmula probabilística fue de 4,94, la cual se aproximó a número entero dando como resultado 5 bases de datos bajo un nivel de confianza de 98%, probabilidad de éxito del 80% y probabilidad de fracaso del 20%.



POBLACIONES FINITAS
(<100.000 INDIVIDUOS)

$$n = \frac{2,33^2 * 0,8 * 0,2 * 5}{0,02^2 * (5-1) + 2,33^2 * 0,8 * 0,2}$$

n= 4,94

Figura 19 Fórmula de poblaciones finitas para el tamaño de la muestra 1
Fuente: Elaboración propia

Sin embargo, no todas las cinco (5) bases de datos cumplían con los criterios de selección establecidos en el muestro teórico, por lo que se reduce el tamaño de unidades muestrales a una (1) base de datos acorde la selección de muestra teórica.

Así en entonces y de acuerdo con lo descrito en el diseño de investigación, la selección de la muestra fue dada de la siguiente manera, basada en la experiencia y conocimiento del tema del investigador en su rol desde la Oficina de TI, apoyados en que como dice Nogales (2004) “La experiencia del investigador y su conocimiento del tema y del colectivo implicado sirven de base para determinar el criterio a seguir en la selección muestral” (p. 154) :

Es así como bajo muestreo teórico se realiza la selección de una de las cinco (5) bases de datos que gestiona la Dirección de Apropiación; Teletrabajo, Ciudadanía Digital, RedVolución, EnTIC Confío y ConverTIC, de acuerdo con los siguientes ítems tenidos en cuenta como constructo teórico elaborado por el autor del presente trabajo, en donde "...se lleva a cabo siguiendo las tipologías o perfiles definidos conceptualmente" que "el investigador teoriza que pueden influir en los significados del fenómeno estudiado" (Silva, 2006, p. 44).

- Volumen de Datos: significa la cantidad de registros reportados en cada base de datos. Se determinará un volumen considerable de 500.000 registros en adelante.
- Cantidad de Devoluciones por no Conformidad de los datos: hace referencia a la cantidad de veces que es devuelta al responsable, una base de datos por presentar datos no conformes.
- Diversidad de Errores: se refiere a la tipología de errores encontrados en las bases de datos por el autor del presente trabajo de grado, dado su conocimiento técnico y experiencia.
- Importancia Misional: se refiere a la frecuencia mensual de solicitudes de información por parte de los directivos, proveniente de dichas bases de datos.

Muestra 2: el área del MinTIC en la que se enfoca el desarrollo del estudio es la Dirección de Apropiación, la cual cuenta con 60 recursos humanos entre funcionarios y contratistas, de los cuales 5 son los responsables directos de la gestión de las bases de datos de cada uno de los proyectos de la Oferta Institucional en esta Dirección, y quienes son tenidos en cuenta como "informantes que posean un conocimiento general amplio sobre el tópico a indagar, o informantes que hayan vivido la experiencia sobre la cual quiere ahondar" (Bonilla-Castro, E. & Rodríguez Sehk, 2005, p. 138).

Para la selección del tamaño de la muestra de los responsables, en primera instancia se aplicó la fórmula de poblaciones finitas, teniendo como universo a los 5 responsables de gestión directos en la Dirección de Apropiación, este cálculo arrojó como resultado 5 unidades muestrales.

POBLACIONES FINITAS
 (<100.000 INDIVIDUOS)

$$n = \frac{2,33^2 * 0,8 * 0,2 * 5}{0,02^2 * (5-1) + 2,33^2 * 0,8 * 0,2}$$

n= 4,94

Figura 20 Fórmula de poblaciones finitas para el tamaño de la muestra 2
 Fuente: Elaboración propia

No obstante, en el desarrollo de la investigación no se logró cubrir el tamaño de la muestra dado que no fue posible acceder en su totalidad, por lo que esto representa una limitación del estudio y se procede entonces a definir como muestra la totalidad de las unidades a las que se tuvo acceso.

Muestra 3: el tamaño de la muestra se determina a partir de la aplicación de la fórmula para poblaciones finitas, en donde el resultado arrojado es de 6 unidades muestrales, teniendo como universo a los 6 expertos identificados que han tenido relación con la Oficina de TI en temas de gestión de datos, bajo un nivel de confianza de 98%, probabilidad de éxito del 80% y probabilidad de fracaso del 20%. A continuación, se relaciona la aplicación de la fórmula:

POBLACIONES FINITAS
 (<100.000 INDIVIDUOS)

$$n = \frac{2,33^2 * 0,8 * 0,2 * 6}{0,02^2 * (6-1) + 2,33^2 * 0,8 * 0,2}$$

n= 5,98

Figura 21 Fórmula de poblaciones finitas para el tamaño de la muestra 3
 Fuente: Elaboración propia

Teniendo en cuenta las 6 unidades muestrales, se procede a aplicar el muestreo teórico de acuerdo con los siguientes requisitos en donde “se pretende que la muestra de participantes esté compuesta por casos seleccionados en función de que posean determinadas características relevantes en relación con los objetivos del estudio” (Quintanilla Cobián, García Gallego, Rodríguez Fernández, Fontes de Gracia, Sarriá Sánchez, 2019, párr. 4) :

- Profesionales en el área de Tecnologías de la Información.
- Profesionales que laboren en la industria de los datos.
- Profesionales con trayectoria mínima de 4 años.

Aunque el resultado probabilístico fue de 6 unidades muestrales y todas las unidades muestrales cumplen con los criterios establecidos, solo se pudo acceder a cuatro (4) finalmente de manera exitosa.

3.1.2 Herramientas para recolección de datos

Matriz de Diagnóstico: la matriz de diagnóstico que permite determinar el estado de calidad de la base de datos seleccionada se diseña de la siguiente manera como estrategia de integración multimétodo y es tomada en cuenta como instrumento de enfoque cuantitativo, por manejar resultados numéricos en su evaluación, pero busca complementar los resultados de los otros instrumentos sin cambiar el enfoque principal. Dicho en otras palabras “En esta estrategia, un enfoque se destaca sobre el otro, de hecho, la investigación se desarrolla desde un enfoque predominante y los resultados obtenidos se fortalecen anexándole un componente de otro enfoque, bien sea en el momento del diseño, la recolección de datos o el análisis” (Durán, 2019a, p. 50):

1. Se deben enlistar de manera vertical los campos de la entidad (tabla) de la base de datos a evaluar.
2. Se deben enlistar las dimensiones seleccionadas de los marcos de referencia del presente trabajo de grado de manera horizontal para el diagnóstico.

Entrevistas a profundidad

Dentro de las herramientas para la recolección de datos en esta fase, se aplican entrevistas a profundidad a los dos tipos de muestra seleccionados (Muestra 2 y 3). Las entrevistas son semiestructuradas y tienen una duración máxima de 30 minutos. Las guías de las entrevistas están estructuradas en dos temas: Guía I) Descripción de la problemática actual en la Dirección de Apropiación frente al tema de gestión de datos y Guía II) Recomendaciones y buenas prácticas en gestión de datos para el planteamiento de la solución. Estas guías están diseñadas a partir de preguntas orientadoras alineadas con los objetivos de la investigación.

Las entrevistas son realizadas de manera personal y en las instalaciones del MinTIC a los responsables de la gestión de datos. Para el segundo grupo muestral se realizan en un lugar definido entre el experto y el autor del presente trabajo de grado. Como herramienta de apoyo para el desarrollo de las entrevistas se utiliza grabación de audio con previa autorización de cada entrevistado.

Las guías de entrevistas se encuentran relacionadas en el apartado de anexos.

3.1.3 Análisis

Fase 1 Matriz de diagnóstico

El análisis de la base de datos seleccionada (Ciudadanía Digital) mediante la matriz de diagnóstico se realiza de la siguiente manera:

1. Por cada dimensión se define un valor booleano (SI / NO), con el cual se identificará si el campo evaluado cumple o no con el criterio.

2. Cuando se asignan los valores (SI / NO) en cada casilla, resultado del análisis de los datos, se realiza un conteo del número de veces de la palabra (SI), es decir cuántos cumplen por registro.

3. Con el fin de poder definir un resultado, se utilizan unos rangos para discriminar si es “Baja”, “Media” o “Alta” la calidad de los datos. Este rango se diseña de la siguiente manera:

- a) Se considerará Alto: Si el conteo de los (SI) es mayor o igual a 6.
- b) Se considerará Medio: Si el conteo de los (SI) está entre 4 y 5.
- c) Se considerará Bajo: Si el conteo de los (SI) es menor o igual a 3.

4. Contar el total de coincidencias de los diferentes niveles de calidad (“Baja”, “Media” o “Alta”) obtenidas para cada uno de los campos evaluados; el nivel que presente el mayor número de conteos determina la calidad de la base de datos analizada.

Fase 2 Entrevistas a profundidad

Para la realización de análisis de dichos datos recolectados en las entrevistas como herramienta de investigación, se debe tener en cuenta que este análisis debe ser sistemático, seguir una secuencia y tener un orden claro (Hernández Sampieri et al., 2010).

1. **Realizar entrevistas:** se obtiene la información necesaria a través de la entrevista como instrumento de investigación aplicado los dos grupos muestrales.

2. **Capturar y transcribir:** para la captura de la información se tienen en cuenta los siguientes medios que permitan guardarla, como lo son grabaciones de audio digital, notas tomadas por parte del entrevistador mediante papel y bolígrafo o digital. Al terminar las sesiones de entrevista, son transcritas a un formato que sea legible.
3. **Ordenar la información:** en este paso se realiza la segmentación mediante bloques temáticos que previamente se han establecido en la organización en las guías de entrevista y que han sido definidos y orientados hacia objetivos. En donde se extraen y ordenan fragmentos de las entrevistas transcritas en una parrilla organizada por los bloques temáticos “correspondientes a cada gran dimensión de análisis” (Seid, 2016).
4. **Codificar la información:** para codificar la información se realiza mediante la agrupación de temas, luego de haberla agrupado se hace una clasificación de estas mismas para formar categorías que permiten concentrar las ideas, conceptos o temas similares descubiertos por el investigador.
5. **Integrar la información:** relacionar las categorías obtenidas en el paso anterior, entre sí. Seguidamente, para la Muestra 2 se producirá una triangulación entre respuestas de la entrevista de la Guía I y una búsqueda en la lista de procedimientos y/o manuales referentes a gestión de datos en el MIG (Modelo Integrado de Gestión) del MinTIC. Para, la Muestra 3 se realizará una triangulación entre las categorías obtenidas y los fundamentos teóricos del presente trabajo de grado.

El proceso de la codificación fragmenta todas las transcripciones hechas por el investigador en categorías separadas ayudan ver cada detalle específico, para determinar que aporta al análisis. Una vez que se han encontrado esos conceptos y temas individuales, se deben relacionar entre sí para poder elaborar una explicación integrada. Luego, el material se compara entre las diferentes categorías, buscando los vínculos que puedan existir entre estas y el marco conceptual.

El procedimiento realizado para el análisis de las entrevistas se describe a continuación:

Se realizaron 6 entrevistas en total. Para la muestra 2 (Responsables de Gestión) se efectuaron tres (3) entrevistas y para la muestra 3 (Expertos), se realizaron cuatro (4) entrevistas, de las cuales ninguna superó los 30 minutos estipulados para su desarrollo y no se presentó dificultad alguna dado que los dos grupos de entrevistados tenían amplio conocimiento sobre los temas tratados en el cuestionario.

Las entrevistas se aplicaron basadas en las guías propuestas y se registró bajo grabaciones de audio previa autorización. Luego de obtener los audios, se

transcribieron las entrevistas pregunta por pregunta, entrevista por entrevista. Una vez transcritas todas las entrevistas se procedió a organizarlas en parrillas segmentadas por bloques temáticos.

De esta manera se organizaron las 3 entrevistas del grupo de la muestra 2 (Responsables de Gestión) y las 4 entrevistas de la muestra 3 (Expertos) adicionalmente dividida por subtemas. Se ilustra este paso con la siguiente tabla:

Nombre Grupo Muestral 1		
BLOQUE TEMÁTICO 1	Subtema 1	Respuesta
	Subtema 2	Respuesta
	Subtema 3	Respuesta
EXPERTO 1		
INTRODUCCIÓN AL PROBLEMA DESDE LA PERSPECTIVA DE LOS EXPERTOS	Problemas en las Organizaciones.	"La inconsistencia de la información" tienes que hacerle tratamiento...para que alguna manera cobre en valor y puedan apoyar veces esa toma decisiones. *la fiabilidad de la información de los datos que contiene".
	Frecuencia de los problemas de datos en las organizaciones	"está en presente ellas. todas son frecuentes... sí, te puedo asegurar que no existe una compañía que tengo una calidad mediana de datos."
	Causas	"no se han tenido en cuenta validaciones, reglas, depuraciones, homologaciones".*En Colombia sobre todo no hay una política real de elegir las implicaciones que contiene el diligenciar mal la información."

Tabla 2 Organización de transcripción de entrevistas por parrillas segmentadas
Fuente: Elaboración Propia

Por cada respuesta de cada entrevistado dividida en los dos grupos muestrales (Responsables de Gestión y Expertos) se resaltaron palabras que más destacaban la idea en general de cada respuesta y se agruparon todas las palabras resaltadas de todos los entrevistados del mismo grupo muestral por tema y se analizó concentrando todas las ideas de cada entrevistado por cada tema. La siguiente tabla ilustra la forma como se desarrolló este ejercicio:

Nombre Grupo Muestral 1					
BLOQUE TEMÁTICO 1		Subtema 1		Res-pues-ta	
		Subtema 2		Respuesta	
		Subtema 3		Respuesta	
INTRODUCCIÓN AL PROBLEMA DESDE LA PERSPECTIVA DE LOS EXPERTOS		EXPERTO 1		EXPERTO 2	
		Problemas en las Organizaciones.	"La inconsistencia de la información " tienes que hacerle tratamiento...para que alguna manera cobre en valor y puedan apoyar veces esa toma decisiones. *" la fiabilidad de la información de los datos que contiene".	Problemas en las Organizaciones.	"La gente no conoce qué datos tiene 1. La gente no conoce cuál es la calidad de sus datos ." " desconocimiento del valor del dato " Por eso es que no se fijan en si los datos son o no de calidad, si los datos son o no consistentes."
		Frecuencia de los problemas de datos en las organizaciones	"está en presente ellas. todas son frecuentes ... sí, te puedo asegurar que no existe una compañía que tengo una calidad mediana de datos."	Frecuencia de los problemas de datos en las organizaciones	"Eso se encuentra mucho en las organizaciones ." "entidades que visito podría estar diciendo que el 90 por ciento de las entidades tienen problemas con la Data".
		Causas	" no se han tenido en cuenta validaciones, reglas, depuraciones, homologaciones ".**"En Colombia sobre todo no hay una política real de elegir las implicaciones que contiene el diligenciar mal la información."	Causas	"porque es que llevan conviviendo con ese problema toda la vida y funcionan." "Entonces pues al final del ejercicio no es algo que les duela , les empieza a doler cuando la alta gerencia pide información y las áreas que operan la data tienen que sentarse a procesarla y cuando se sientan a procesarlos se dan cuenta que es data incompleta que no tienen información consistente que no es Data segura que no es data

Respuestas principales Expertos/ Bloque Temático 1

-La **inconsistencia de la información** **la fiabilidad**

-**presente frecuentes**

-**validaciones, reglas, depuraciones, homologaciones no hay política real**

-**no conoce la calidad de sus datos desconocimiento del valor del dato**

- **mucho en las organizaciones**

-**llevan conviviendo no es algo que les duela empieza a doler tienen que sentarse a procesarla se dan cuenta que es data incompleta consistente no es data confiable"**

Análisis Bloque temático 1: Se concluye que... para el bloque temático 1.

Análisis Bloque temático 1: El problema que se presenta en muchas organizaciones frecuentemente es la inconsistencia de la información, esto debido al desconocimiento del estado o calidad que presentan los datos que se están recibiendo de las diversas fuentes de las que se obtienen. A su vez, esto es producido por la falta de políticas, reglas, lineamientos, procesos que permitan ejercer control sobre la gestión del dato para que, desde su captura, la cual por lo general es manipulado por personas, se empiece a dar valor al dato.

Tabla 3 Concentración de ideas por bloque temático

Fuente: Elaboración propia

De acuerdo con los conceptos o ideas similares resaltadas por el investigador, se realizaron análisis de las respuestas de los entrevistados destacando conceptos relevantes entre sí y encontrando coincidencias de un entrevistado a otro en sus respuestas. Se aplicó la triangulación como herramienta de la siguiente manera para la muestra 2 (Responsables de Gestión) entre las ideas concentradas puntualmente dentro del bloque temático de “*Procedimientos actuales*” y la revisión en el MIG.

Bloque temático 2 (Responsables de Gestión)	Hallazgos en el MIG sobre el Bloque temático 2	Triangulación
Análisis de las respuestas de los Responsables de Gestión Muestra 2 sobre el Bloque Temático.	Revisión en el MIG sobre el Bloque Temático 2.	Enlace de lo que dicen los Responsables de Gestión VS lo que se encuentra en el MIG.
Instrumentos, Metodologías Para La Gestión De La Calidad De Datos Utilizados Por Los (Responsables De Gestión)	Instrumentos, Metodologías Para La Gestión De La Calidad De Datos regidos por el MIG	Triangulación
Los responsables de gestión de los datos de la Dirección de Apropiación del MinTic afirman no contar con ninguna estrategia para la gestión de calidad de los datos que manipulan en sus actividades del día a día.	En la revisión de manuales, planes, procedimientos y guías sobre calidad de datos no se encuentran instrumentos para las actividades de gestión calidad de datos.	No se encontró ninguna documentación asociada a, planes, procedimientos, manuales o guías relacionadas con calidad de datos. De lo que se infiere, la afirmación de los responsables es correcta y no tienen a la mano un referente para realizar sus actividades dentro de la Dirección de Apropiación correspondientes a la gestión de los datos con calidad.

Tabla 4 Triangulación Bloque Temático 2 Y MIG
Fuente: elaboración propia

Para la muestra 3 (Expertos) se hizo triangulación entre los conceptos o ideas similares encontradas por el investigador en las respuestas de los bloques temáticos “*Entorno ideal para la gestión de los datos y marcos de referencia*” contra el Marco Conceptual.

CRITERIOS DE CALIDAD DE DATOS Y MARCO QUE RECOMIENDA	MARCO CONCEPTUAL	TRIANGULACIÓN
<ul style="list-style-type: none"> Identificar los datos que son importantes para la entidad. Identificar procesos. Definir actividades, funciones, roles y responsabilidades. Habilitar las capacidades técnicas y conceptuales de la labor que realizan. 	<p>Mike 2.0: los conocedores de los datos, que han sido seleccionados como de alto impacto.</p> <p>Según (Thomas, s.f.) en un artículo del Data Governance Institute, el gobierno de datos se define como un sistema de derechos y responsabilidades de decisión para procesos relacionados con la información y se ejecutan con base en los modelos diseñados previamente en los que se describe quienes pueden tomar determinadas medidas, sobre qué información, cuándo, y bajo qué métodos y/o herramientas.</p> <p>el DAMA definen que un modelo de gobierno de datos que respalde la estrategia de negocio mínimo debe incluir:</p> <ul style="list-style-type: none"> •Alineación estratégica: impulsores, misión y visión. •Objetivos, principios y políticas. •Procesos relacionados con la operación de los datos. •Reglas de negocio. •Roles y responsabilidades. <p>Según DAMA: identificación de datos críticos y reglas de negocio “...programa de gobierno de datos y se cuente con profesionales capacitados para ello”</p>	<p>Se puede evidenciar con total claridad que lo que presenta el marco conceptual bajo las definiciones de criterios relevantes y buenas, concuerdan con los criterios y buenas prácticas que los expertos entrevistados recomiendan como clave a tener en cuenta para adaptar y adoptar en una estrategia de calidad de datos.</p> <p>Según Mike 2.0, se debe realizar una identificación o priorización de los datos a los que se les va a realizar seguimiento para la mejora de la calidad.</p> <p>Según (Thomas, s.f.) el gobierno de datos define los derechos y responsabilidades de decisión para procesos relacionados con la información, esto implica que debe realizarse una identificación de procesos de información. El DAMA recomienda que cualquier proyecto de calidad, debe estar enmarcado en un gobierno de datos, que incluya entre otros: roles, responsabilidades para las personas y que permita habilitar a las personas involucradas para ejercer dichas actividades.</p>

Tabla 5 Triangulación bloque temático 2 y respuestas de expertos.

Fuente: elaboración propia

3.2 Paso a paso de la metodología

Objetivo Específico	N°	Actividades	Fuentes	Herramienta	Resultado
Realizar un diagnóstico de la situación actual en la Dirección de Apropiación,	1	Seleccionar muestra 1 (base de datos) para aplicar la matriz de diagnóstico.	Bases de datos de la Dirección de Apropiación.	Lista de ítems establecidos como constructo teórico para seleccionar la muestra.	Base de datos seleccionada.

Objetivo Específico	N°	Actividades	Fuentes	Herramienta	Resultado
respecto a la gestión de datos.		Seleccionar muestra 2 (responsables de gestión de datos) para aplicar entrevistas.	Población responsable de la gestión de datos en la Dirección de Apropiación.	Muestreo por conveniencia.	Responsables seleccionados.
	2	Aplicar matriz de diagnóstico.	Muestra 1 (Base de datos seleccionada).	Matriz de Diagnóstico.	Resultado de la evaluación del nivel de calidad de la base de datos.
	3	Aplicar entrevista.	Muestra 2 (responsables de gestión de datos seleccionados).	Guía I de entrevista.	Guía I aplicada.
	4	Realizar el Diagnóstico del estado actual de la gestión de datos de la Dirección de Apropiación.	Resultado de la matriz de diagnóstico. Análisis de respuestas de las entrevistas de la Guía I (responsables de la gestión de datos) .	Codificación e Integración de los resultados obtenidos.	Diagnóstico situacional de la gestión de datos en la Dirección de Apropiación.
Evaluar y seleccionar criterios relevantes y buenas prácticas de los marcos de referencia mencionados en el presente trabajo, para adaptarlos en un plan que atienda las necesidades de gestión de datos de la oferta institucional en la Dirección de Apropiación.	5	Seleccionar muestra 3 (expertos) para aplicar entrevistas.	Población de expertos en Calidad de Datos.	Lista de requisitos establecidos como constructo teórico para seleccionar la muestra.	Profesionales expertos seleccionados.
	6	Aplicar entrevista.	Muestra 3 (expertos seleccionados).	Guía II de entrevista.	Guía II aplicada.
	7	Analizar entrevistas aplicadas.	Respuestas de las entrevistas de la Guía II (expertos).	Codificación e Integración de los resultados obtenidos.	Análisis de las entrevistas.
	8	Identificar y seleccionar criterios y buenas prácticas.	Marco conceptual y análisis de entrevistas a expertos.	Integración de los resultados obtenidos. (Triangulación)	Criterios seleccionados para adaptarlos al plan de gestión de datos.

Objetivo Específico	N°	Actividades	Fuentes	Herramienta	Resultado
Proponer las acciones de mejora a los procedimientos actuales de gestión de datos de la Dirección de Apropiación.	9	Identificar los procedimientos actuales de gestión de datos de la Dirección de Apropiación.	Respuestas a las entrevistas de la Guía I (responsables de gestión de datos) y revisión de procedimientos en el MIG (Modelo Integrado de Gestión).	Codificación e Integración de los resultados obtenidos.	Identificación de los procedimientos de gestión de datos en la Dirección de Apropiación.
	10	Proponer acciones de mejora.	Criterios relevantes seleccionados. (resultado paso 6). Procedimientos de gestión de datos identificados.	Integración de los resultados obtenidos. (Triangulación)	Propuesta de acciones de mejora para construir el plan de gestión.
Identificar los beneficios de proponer un plan de gestión de calidad de datos para la Dirección de Apropiación.	11	Revisar entrevistas.	Respuestas del bloque temático: <u>“Consideraciones De Proponer Un Plan De Gestión”</u> de las entrevistas de la Guía I (responsables de gestión de datos) y respuestas del bloque temático <u>“El Plan De Gestión De Calidad De Datos Como Herramienta De Solución.”</u> de las entrevistas de la Guía II (expertos).	Codificación e Integración de los resultados obtenidos.	Identificación de posibles beneficios del plan de gestión a proponer.
Proponer un plan de gestión de calidad basado en las acciones de mejora enlistadas y los criterios	12	Revisar entrevistas.	Respuestas del bloque temático: <u>“El Plan De Gestión De Calidad De Datos”</u>	Codificación e Integración de los resultados obtenidos.	Consideraciones para tener en cuenta en la propuesta del Plan de Gestión

Objetivo Específico	N°	Actividades	Fuentes	Herramienta	Resultado
seleccionados de los marcos de referencia, que permita en el corto, mediano y largo plazo cerrar brechas y obtener mejoras en los datos de la oferta institucional de la Dirección de Apropiación.			<u>Como Herramienta De Solución</u> de las entrevistas de la Guía II (expertos).		de Calidad de Datos a desarrollar en el capítulo 4.

Tabla 6 Resumen de la Metodología
Fuente: Elaboración propia

3.3 Fuentes de información: personas, instrumentos

Para la *muestra 1* se realiza una selección entre cinco (5) bases de datos que actualmente reporta la Dirección de Apropiación; Teletrabajo, Ciudadanía Digital, RedVolución, EnTIC Confío y ConverTIC.

De acuerdo con los ítems establecidos como constructo teórico, la base de datos seleccionada para aplicar la Matriz de Diagnóstico es la de Ciudadanía Digital, la cual presenta un volumen de datos entre seiscientos cincuenta mil (650.000) y un millón (1'000.0000) de registros. Esta base de datos ha sido devuelta al responsable hasta cinco (5) veces en cada periodo de reporte por la no conformidad de los datos. Los errores que presenta son: falta de integridad, falta de oportunidad, falta de consistencia y falta de completitud en los datos. La importancia misional de esta base de datos está dada por su alto nivel de solicitud respecto las demás bases, por parte de los directivos de la Entidad y la ciudadanía que demanda la información actualizada.

No	Proyecto	Volúmen de Datos (No de Registros)	No de Devoluciones por no conformidad	Diversidad de Errores	Importancia Misional (Solicitudes Por Mes)
1	Teletrabajo	60	2	* Falta de Integridad * Falta de Oportunidad * Falta de Consistencia	2
2	Ciudadanía Digital	720.000	5	* Falta de Integridad * Falta de Oportunidad * Falta de Consistencia * Falta de Completitud	3
3	En Tic Confío	12.000	1	* Falta de Integridad * Falta de Completitud	1
4	ConverTIC	46.000	2	* Falta de Oportunidad * Falta de Validez	2
5	Redvolución	900	2	* Falta de Completitud	1

Tabla 7 Criterios teóricos de selección de la muestra 1
Fuente: Elaboración propia

Para la *muestra 2*, se establece que el tamaño de la muestra escogida es de tres (3) de los cinco (5) responsables de la gestión de datos en la Dirección de Apropiación, debido a que solo esas (3) fueron las unidades muestrales que se lograron acceder de manera exitosa. Por motivos de tiempos entre el investigador y los otros dos (2) responsables de gestión no se logró efectuar la aplicación del instrumento de la entrevista, por lo que limitó el tamaño de las unidades muestrales.

El tamaño de la *muestra 3*, se determina ubicando el perfil de expertos que cumple con los ítems de selección establecidos. Dicha labor se realiza a partir de la identificación de profesionales de las empresas que han trabajado en proyectos de TI con el Ministerio para luego contactarlos e invitarlos a participar en el estudio.

Se identifican 6 empresas y se contacta un profesional experto por cada una de ellas. Es decir, se obtienen seis (6) profesionales de los cuales cuatro (4) acceden a la realización de una entrevista. Por motivos ajenos al investigador, (2) de los expertos no logran ser concretados para aplicar la entrevista.

A continuación, se relaciona la tabla con los diferentes criterios tenidos en cuenta:

No.	Empresa/ Entidad	Profesionales en el área de Tecnologías de la Información.	Profesionales que laboren en la industria de los datos.	Profesionales con trayectoria mínima de 4 años.	Acepta Entrevista	Rol
1	Business Insights	X	X	X	X	Gerente de Servicios
2	Controles Empresariales	X	X	X	X	Arquitecto de Datos
3	CRC	X	X	X	X	Líder de Gobierno
4	Gati Consultores	X	X	X	X	Arquitecto Empresarial
5	ldata	X	X	X	-	Arquitecto de Datos
6	Informese	X	X	X	-	Presidente

Tabla 8 Criterios teóricos de Selección de la Muestra 3
Fuente: Elaboración propia

Los cuatro recursos humanos seleccionados son Profesionales en el área de Tecnologías de la Información y trabajan actualmente en temas relacionados con la industria de datos para Entidades Públicas, Privadas y Mixtas. Todos los profesionales cuentan con trayectoria de más de 4 años de experiencia profesional y presentan los siguientes roles actualmente:

1. Gerente de Servicios / Business Insights.
2. Arquitecto de Datos/ Controles Empresariales.
3. Líder de Gobierno de Análisis de Datos/ CRC (Comisión de Regulación de Comunicaciones).
4. Arquitecto Empresarial / Gati Consultores.

En cuanto a la estructura de las guías de entrevista I y II, se desarrollaron bajo preguntas orientadoras dirigidas hacia la consecución de los objetivos específicos, de acuerdo con el siguiente cuadro se obtuvieron los cuestionarios que se relacionan en los Anexos 1 y 2.

Objetivo	Pregunta
Realizar un diagnóstico de la situación actual en la Dirección de Apropiación, respecto a la gestión de datos.	<p>¿Conoce cuáles son los criterios mínimos de calidad para los datos?</p> <p>¿Realiza una revisión de la información que le entregan los operadores?</p> <p>¿Durante el proceso contractual, se exigen las condiciones mínimas para el reporte de información al Ministerio?</p>

Objetivo	Pregunta
	<p>¿Cuentan con un procedimiento, una guía o alguna metodología que le permita realizar control sobre los datos?</p>
<p>Evaluar criterios relevantes y seleccionar buenas prácticas de los marcos de referencia mencionados en el presente trabajo, para adaptarlos en un plan que atienda las necesidades de gestión de datos de la oferta institucional en la Dirección de Apropiación.</p>	<p>¿Tiene conocimiento acerca de lo que es el DAMA?</p> <p>¿Cuáles marcos de referencia considera que se deben tener en cuenta para la implementación de un plan de gestión de calidad de datos??</p> <p>Dado que las ISO, son estándares internacionales y el DAMA es reconocido internacionalmente también, ¿cuál considera que es un mejor referente para ser adaptada a las necesidades de la entidad?</p>
<p>Proponer las acciones de mejora a los procedimientos actuales de gestión de datos de la Dirección de Apropiación.</p>	<p>¿Cuáles criterios de control de calidad considera que deben tener en cuenta las áreas de una organización, cuyas fuentes no provienen directamente de sistemas de información?</p> <p>¿Considera, que, para los procesos de gestión de información, las personas encargadas, deben tener conocimientos acerca de la labor que realizan?, qué tipo de conocimientos?</p> <p>¿Que deben implementar las áreas de una organización para mejorar la gestión de sus datos?</p>
<p>Identificar los beneficios de proponer un plan de gestión de calidad de datos para la Dirección de Apropiación.</p>	<p>¿Cree que la implementación de un plan de calidad representa un costo alto para la entidad?</p> <p>¿Considera que la adopción de un plan de gestión, le permitirá mejorar los tiempos de generación y reporte de información?</p> <p>¿Cuál es la razón principal por la que las entidades deben implementar planes de calidad de datos?</p> <p>¿Como se ven reflejados los beneficios de la adopción de un plan de calidad, de cara a los involucrados en los procesos de información?</p>

Objetivo	Pregunta
<p>Proponer un plan de gestión de calidad basado en las acciones de mejora enlistadas y los criterios seleccionados de los marcos de referencia, que permita en el corto, mediano y largo plazo cerrar brechas y obtener mejoras en los datos de la oferta institucional de la Dirección de Apropiación.</p>	<p>¿considera que es necesario, implementar o adoptar alguna estrategia que les permita mejorar sus procesos de reporte?</p> <p>¿Si se adopta un plan de gestión en el área, cree que mejorarían los procesos de reporte de información?</p> <p>¿por qué considera que la implementación de un plan de gestión de calidad de datos mejoraría notablemente los procesos de reporte de información en las áreas?</p> <p>¿Por qué es importante contar con una metodología o plan o algún otro tipo de instrumento para la gestión de los datos en las organizaciones?</p>

Tabla 9 Tabla de preguntas orientadoras

Fuente: Elaboración propia

4 Resultados de la investigación

En este capítulo se describirá el desarrollo de la aplicación del diseño metodológico y sus resultados, de acuerdo con cada objetivo propuesto:

4.1 Diagnóstico situacional en la Dirección de Apropiación, respecto a la gestión de datos

En la presente sección se realizará un diagnóstico de la situación actual en la Dirección de Apropiación, respecto a la gestión de datos.

Con base en las recomendaciones de los marcos de referencia y las buenas prácticas mencionadas, se realiza el perfilamiento de los datos de la base de datos de “Ciudadanía Digital” a continuación:

Perfilamiento de los datos

- Campos candidatos para llave primaria

Teniendo en cuenta que el ciudadano puede realizar una certificación por cada curso se proponen los siguientes campos para que conformen la llave primaria de la tabla.

- Certificaciones
- Tipo_Documento
- Documento

Aunque actualmente no se encuentran registros duplicados, es un riesgo latente que se puede llegar a presentar en un futuro; por consiguiente, se propone que se utilicen los tres campos listados anteriormente como llave primaria de la tabla; así se logra evitar duplicidad de datos.

- Distribución de longitud por columnas

Columna *	Longitud mínima	Longitud máxima
Tipo_Documento	4	23
Documento	6	11
Nombres	3	37
Apellidos	1	32
Email	9	52

Telefono	1	12
Departamento_Actual	1	24
Ciudad_Actual	1	36
Direccion	1	110
Nivel_Educativo	1	33
Estrato	1	1
Fecha_de_nacimiento	1	31
Genero	4	9
Discapacidad	1	12
Ocupacion	1	96
Fecha_Registro	0	16
Comunidad_Etnica	0	17
COD_DANE_MUN	0	5
Codigo Departamento Actual	0	2

Tabla 10 Perfilamiento de longitud de campos
Fuente: Elaboración propia

* Las columnas marcadas con asterisco (*) son nombres de campos, no incluyen ortografía.

Obteniendo esta distribución se identifican campos con valores que no corresponden; un ejemplo es observar la longitud de los caracteres que se digitan en el campo “Telefono” donde claramente se evidencia error debido a que los números de teléfono no pueden ser de un solo dígito.

En el análisis ejecutado se evidencia que de los 22.409 registros que contiene la tabla, 11.806 registros contienen un número de teléfono con un valor de un (1) solo dígito; es decir que el 52,68% de los números de teléfonos se encuentran mal registrados.

- Columnas con valor NULL

Columna *	Conteo de nulos
Tipo_Documento	0
Documento	0
Nombres	0
Apellidos	0
Email	0
Telefono	0
Departamento_Actual	0
Ciudad_Actual	0
Direccion	0
Nivel_Educativo	0
Estrato	0
Fecha_de_nacimiento	0

Genero	0
Discapacidad	0
Ocupacion	0
Fecha_Registro	5
Comunidad_Etnica	5
COD_DANE_MUN	5
Codigo Departamento Actual	5

Tabla 11 Perfilamiento de campos nulos

Fuente: Elaboración propia

* Las columnas marcadas con asterisco (*) son nombres de campos, no incluyen ortografía.

Solo en algunos campos se encontraron registros Nulos, pero las propiedades de un archivo en formato Excel, permite pasar valores vacíos o Nulos, lo cual mantendrá el riesgo de presentar error en los datos.

- Tipos de dato/Formato

Los tipos de dato digitados en los campos “Email”, “Fecha_de_nacimiento”, “Discapacidad” y “Ocupacion” no concuerdan con el tipo de dato esperado, generando inconvenientes a la hora de manipulación de los datos y/o errores en cálculos de indicadores.

Algunos errores asociados a los tipos de datos se muestran en la siguiente tabla:

Columna *	Tipo de dato esperado	Formato esperado	Valores digitados en base de datos
Email	Varchar	<ul style="list-style-type: none"> • Formato tipo correo electrónico el cual debe contar con el símbolo de “@” y seguido de un dominio válido. 	<ul style="list-style-type: none"> • edgarhumberto12345yaho..es • 3127746768 • Federicopachecomoreno • rojasmariadelcarmen56gmail.com • cl 123 46 45, ca primer piso • 3127746768
Fecha_de_nacimiento	Date	<ul style="list-style-type: none"> • El valor esperado debe cumplir con el formato DD/MM/AAAA. Donde DD = Día, MM = Mes y AAAA = Año 	<ul style="list-style-type: none"> • MiÃ©rcoles, Mayo 11, 1966 • - • 0/01/1900 0:00
Discapacidad	Varchar	<ul style="list-style-type: none"> • Valores con formato tipo texto y que correspondan a una discapacidad. Ejemplo: Auditiva, Cognitiva, Ninguna 	<ul style="list-style-type: none"> • - • 0
Ocupacion	Varchar	<ul style="list-style-type: none"> • El formato debe corresponder a texto y 	<ul style="list-style-type: none"> • - • 3213485150

		debe representar de forma válida una ocupación. Ejemplo: Docente, Pensionado, Médico, Estudiante.	• 3008380047
--	--	--	--------------

Tabla 12 Perfilamiento de tipo de datos y formato
Fuente: Elaboración propia

Matriz de diagnóstico

Ahora bien, acorde con el numeral “3.3 Fuentes de información: personas, instrumentos”. se aplicó el instrumento de evaluación denominado Matriz de Diagnóstico a la base de datos de Ciudadanía Digital, que es una base que contiene el registro de datos de todas las personas que han sido beneficiadas a través de este programa, en la matriz se encontró lo siguiente:

ÁREA:	DIRECCIÓN DE APROPIACIÓN DE TECNOLOGÍAS DE INFORMACIÓN - MINTIC								
BASE DE DATOS:	CIUDADANÍA DIGITAL								
DIMENSIONES DE CALIDAD DE DATOS									
DATO (CAMPO)	EXACTITUD	COMPLETITUD	CONSISTENCIA	INTEGRIDAD	UNICIDAD	VALIDEZ	OPORTUNIDAD	CANTIDAD SI	CALIFICACIÓN
Tipo_Documento	NO	SI	NO	NO	NA	NO	NO	1	Bajo
Documento	SI	SI	SI	SI	SI	SI	NO	6	Alta
Nombres	NO	SI	SI	NO	SI	NO	NO	3	Bajo
Apellidos	NO	SI	SI	NO	SI	NO	NO	3	Bajo
Email	SI	SI	SI	SI	NA	NO	NO	4	Medio
Telefono	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Departamento_Actual	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Ciudad_Actual	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Direccion	NO	SI	NO	NO	NA	NO	NO	1	Bajo
Nivel_Educativo	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Sector	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Estrato	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Fecha_de_nacimiento	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Genero	SI	SI	SI	SI	NA	SI	NO	5	Medio
Discapacidad	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Ocupacion	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Fecha_Registro	SI	SI	SI	SI	NA	SI	NO	5	Medio
Comunidad_Etnica	NO	NO	NO	NO	NA	NO	NO	0	Bajo
COD_DANE_MUN	NO	NO	NO	NO	NA	NO	NO	0	Bajo
Codigo_Departamento_Actual	NO	NO	NO	NO	NA	NO	NO	0	Bajo

Tabla 13 Matriz de diagnóstico aplicada
Fuente: Elaboración propia

En la matriz, se observa el listado de campos que fueron objeto de evaluación de la base de datos y el valor asignado frente a cada una de las dimensiones evaluadas. Al final de cada registro y de acuerdo con el conteo de (SI) el cual indica el cumplimiento, se asigna una calificación descriptiva del nivel de calidad de cada registro de acuerdo con los criterios establecidos en el diseño del instrumento numeral 3.1.3. En la matriz también se puede analizar, que en esta calificación

predomina el resultado de “Bajo”; confirmando de esta manera que los problemas en los datos están asociados a la falta de integridad, de oportunidad, inconsistencia en los tipos de datos relacionados para cada campo y de campos vacíos, es decir la no completitud de los datos.

Acorde con el conjunto de datos entregado por la Dirección de Apropiación, las dimensiones que fueron evaluadas para el diagnóstico junto con los hallazgos se describen a continuación:

- **Exactitud:** los datos registrados en campos como: tipo_documento, cod_dane_mun, fecha_de_nacimiento, ocupación entre otros evaluados, no representan de manera correcta los hechos de la vida real que estos significan para una persona, por ejemplo, tipos de documento denominados como “otro” o fechas de nacimiento y códigos DANE registrados con un “guion (-)” no permiten identificar la edad promedio ni la ubicación de las personas beneficiadas.

Para que un dato sea exacto debe cumplir con las siguientes características:

- a. El dato debe ser preciso o correcto para que pueda ser utilizado y brinde valor a los reportes que se requieran.
- b. El dato puede ser verificado contra otras fuentes de datos.
- c. El dato debe estar actualizado a su última versión.

Algunos datos que presentan problemas de Exactitud son:

Columna *	Descripción
COD_DANE_MUN	Valores como “0” y “-” que se han registrado en esta columna no permiten comprobar el municipio al que pertenece el beneficiario. Estos datos registrados no son correctos ya que al cruzarlos con los código DANE, se evidencia que no existen municipios con estos valores.
Fecha_de_nacimiento	Existen valores “-” los cuales no representan una fecha de nacimiento real; también se concluye al ver este dato (-) que la fecha de nacimiento del usuario no se encuentra actualizada.

Tabla 14 Campos con problemas de exactitud

Fuente: Elaboración propia

- **Completitud:** se detectaron datos vacíos en campos relevantes como, departamento_actual, estrato, nivel educativo, cod_dane_mun, comunidad étnica, ocupación entre otros evaluados. Para MinTIC es importante poder identificar al tipo de población que impacta con estos programas, la ausencia de información impide que se puedan tomar decisiones frente a cuáles son las poblaciones más vulnerables que se están alcanzando y que población falta por llegar.

- **Consistencia:** en la evaluación de esta dimensión, se encontraron varias inconsistencias en campos registrados como: Telefono, Departamento_Actual, Direccion, Fecha_de_nacimiento. Esta consistencia se refiere a contar con datos que no presenten conflicto a la hora de realizar segmentación de información y cualquier otro tipo de cálculos que se requieran para expresar la información.

Características que ayudan a identificar valores que no cumplen con la dimensión consistencia son:

- a. Los valores que se salen del rango promedio.
- b. Valores que representan conflicto a la hora de realizar agregaciones.
- c. Los valores que tienen formato correcto, pero rompen con las reglas del negocio.

A continuación, se muestran algunos datos que no cumplen con esta dimensión:

Columna *	Descripción
Fecha_de_nacimiento	Se evidencian registros con fechas de nacimiento de 1.918; esto indica que existen personas con 101 años que se están certificando en algún curso orientado a las tecnologías de la información. Datos clasificados como atípicos como el anterior, hacen que los promedios de edad sean errados en informes.
Teléfono	En la validación de este campo se evidencia que existen valores que poseen un formato de teléfono fijo correcto, cumpliendo con 7 dígitos y valores numéricos, pero algunos se encuentran con signo negativo; ejemplo de esto es un número registrado de la siguiente manera “-5156171”.

Tabla 15 Campos con problemas de consistencia
Fuente: Elaboración propia

- **Integridad:** se encontró que campos como tipo_documento, cod_dan_mun, departamento_actual entre otros evaluados, no guardan integridad, lo que impide que en algún ejercicio de análisis de información con otras fuentes, se puedan realizar cruces entre bases de datos por la falta de integridad de estos campos, es usual realizar cruces por campos que representan valores únicos y estandarizados, por lo cual si se quisiera realizar cruces por números de documento de beneficiados, no se obtendrían resultados óptimos por los valores que registra el campo tipo documento “CÃ©dula de ciudadanÃ-a” o “CÈdula de ciudadanÌa” u “OTRO”, o si bien se necesitaran realizar cruces

entre bases por las ciudades o municipios en los que han sido beneficiados, no se podrían realizar por valores como “-”, “0” “vacíos”.

- **Unicidad:** los datos almacenados no presentan problema de duplicidad, pero la tabla al no poseer identificador único como una llave primaria, podría llegar a presentar a futuro, inconvenientes de duplicidad.
- **Validez:** los datos con formatos y longitud asociados a los campos “email”, “Telefono”, “Direccion” no cuentan con validez en la información almacenada, por ejemplo, correos registrados con dominios incorrectos, como “gmaiii” o “gmaill.fr”, números de teléfonos con un dígito o 4 dígitos, a todas luces no corresponde a un número de celular o teléfono fijo acorde con el estándar.

Características para que un dato cumpla con la dimensión validez:

- a. El dato debe tener el formato correcto (donde se espera un tipo de dato numérico no digitar texto y viceversa).
- b. El dato debe seguir un patrón o estándar establecido (correo con dominios válidos, números de telefonía móvil de 10 dígitos y con prefijos válidos).

Columna *	Descripción
Email	En la revisión se visualizan datos como “3042013358”, edgarhumberto12345yaho..es, mayracatic18, los cuales no cuentan con el formato correcto que representa una cuenta de correo electrónico válido. Una cuenta de correo debe contar con el símbolo @ y seguidamente debe contar con un dominio válido (Gmail, Yahoo, Hotmail...)
Teléfono	En este campo algunos valores no cumplen con la cantidad de dígitos válidos para un número telefónico; números de teléfono registrados erróneamente son: <ul style="list-style-type: none"> • 573000000000 • 0 Esto impide que la Dirección de Apropiación logre un contacto exitoso con los beneficiarios en los casos que se requieran acercamientos en el proceso.

Tabla 16 Campos con problemas de Validez

Fuente: Elaboración propia

- **Oportunidad:** ninguno de los campos evaluados, cumple con esta dimensión, dado que el conjunto de los campos es entregado en la base de datos y de acuerdo con lo mencionado en la problemática, esta no ha sido oportuna, por lo dispendioso del proceso manual, ninguno de estos campos es accesible cuando se requiere.

Luego de realizar el análisis por cada campo de acuerdo con las dimensiones, se totalizan los resultados de la calificación a fin de conocer el porcentaje equivalente de toda la base de datos de la siguiente manera:

CALIDAD	CONTEO	PORCENTAJE
Alta	1	5%
Medio	3	15%
Bajo	16	80%
TOTAL	20	100%

Tabla 17 Resultados de la evaluación diagnóstica

Fuente: Elaboración propia

La matriz de diagnóstico usada como herramienta cuantitativa dentro del enfoque cualitativo de la investigación, como estrategia de integración multimétodo (Durán, 2019b), aporta en la identificación del estado actual de la base de datos desde el punto de vista técnico, ubicándola acorde con su evaluación en nivel BAJO, dado que el 80% de los campos evaluados no cumplen con las dimensiones mínimas de calidad de datos. Estos inconvenientes pueden llevar a la entidad a seguir presentando demoras y/o errores en los cálculos de indicadores para la toma efectiva de sus decisiones.

Ahora bien, de acuerdo con el análisis de las respuestas a las entrevistas de los responsables de la gestión de datos en la Dirección de Apropiación, se logra identificar una problemática en común para cada programa misional de esta área, los responsables manifiestan dificultades con los operadores por la no entrega a tiempo de las bases y por la estructura de los datos, por ejemplo, el entrevistado 1 y 2 coinciden en que no poseen un formato en común acuerdo con el operador lo que provoca demoras en la entrega de la información al Ministerio.

En cuanto a la consistencia de los datos que reciben, los entrevistados 1 y 3 manifiestan presentar problemas con lo que entregan los operadores, toda vez que al momento contactar a las personas beneficiadas de determinado programa, se torna imposible localizarlas, ya que los datos están incompletos o los números de contacto presentan más o menos cifras de un número telefónico estándar ya sea celular o fijo, de igual manera con el correo registrado, quedándose sin opciones de corroborar la información para poder efectuar los pagos correspondientes al operador. Esto genera que deban solicitar otro tipo de soportes o evidencias para tratar de validar el cumplimiento. El entrevistado 3 también coincide con la falta de consistencia en los datos pero desde otro punto de vista, es decir éste declara que cuando los datos son recibidos y luego de ser procesados para generar reportes

que son solicitados por la alta gerencia, éstos son insuficientes, dado que lo que entrega el operador actual, no concuerda ni es consistente con lo entregado por otros operadores anteriores, lo que impide la estandarización de la información y el levantamiento de históricos bajo las mismas métricas. El entrevistado 2 manifiesta que estas especificaciones no quedan definidas tan claramente en los contratos, por lo que deben empezar a hilar datos de otras fuentes y en ocasiones reportan la misma información que no representa datos reales.

De acuerdo con lo que manifiestan los responsables de la gestión de datos, contractualmente lo único que se acuerda es un formato básico que consideran suficiente al momento de iniciar el proyecto; sin embargo, no hay exigencias frente a la calidad de estos campos y su proceso de captura. El entrevistado 3 por ejemplo, manifiesta que el operador realiza de manera manual y en formatos físicos el proceso de captura, por lo que no existe la aplicación de reglas mínimas de negocio ni control sobre los datos capturados y esto, sumado a que en muchas ocasiones las demandas de información por parte de los diferentes interesados no se encuentran contempladas en el formato inicial, conlleva a una dificultad en la generación de reportes efectivos que suplan las necesidades de los consumidores, esto sucede porque no existe una sinergia y comunicación constante entre las diferentes áreas que les permita interactuar de manera clara y procedimental para el reporte de información, de manera tal que puedan alinear las necesidades de información de los interesados, con los datos reales que produce la ejecución de los proyectos.

A todo esto, los responsables indican que el proceso de revisión que realizan es manual y se limitan únicamente a la cantidad de registros que reportan en la base, por lo que usualmente si esperan 10.000 beneficiados, lo que se valida en la base, es que existan 10.000 registros, este es el criterio de aceptación del Ministerio de cara al operador. Así pues, los entrevistados 2 y 3 responden no conocer ningún proceso o procedimiento sobre el cual se puedan basar para realizar toda o parte de la gestión del ciclo de vida de los datos de la Dirección, y el entrevistado 3 destaca enfáticamente que dichas actividades las deben realizar manualmente, tomándole gran parte del tiempo por los reprocesos que deben realizar para lograr una data conforme.

En este punto del análisis se han identificado algunas de las causas de la problemática que presenta la Dirección, no obstante, se logra identificar una causa más que se suma a todo lo anterior. Cuando se les preguntó a los responsables de la gestión de datos de la Dirección si tenían conocimientos en conceptos de calidad de datos, respondieron que no de manera técnica pero que por el ejercicio de sus actividades y lo que les demanda, tienen un conocimiento empírico de lo que se trata y de cómo validarlo, el entrevistado 3 manifiesta que incluso había diseñado en algún momento, una matriz de campos estándar que denominó “enfoque

diferencial” a fin de identificar la población beneficiada de los proyectos, pero que no ha sido tomada en cuenta para establecer a todos los proyectos ni para involucrarla dentro del proceso contractual de la Dirección, por lo que únicamente la usa ella para sus fines de información. Esto lo que genera es que exista una desalineación de datos que se capturan en todos los proyectos de cara a los interesados.

Dado la causa anterior, los responsables no hacen uso de ningún método o herramienta que les permita realizar una medición y mejora continua a la calidad de sus bases de datos, puesto que no conocen ni procedimientos internos ni estándares generales que tomen como referencia para la gestión de datos que les permita optimizar las actividades que ejecutan y mejorar los niveles de calidad de sus fuentes.

A los expertos entrevistados, también se les indagó acerca de las posibles causas que a través de su experiencia ellos han podido detectar en las diferentes compañías en donde han laborado y que son semejantes a la situación que vive la Dirección de Apropiación. El experto entrevistado 3 manifiesta que una de las causas más comunes en las empresas, se encuentra en los procesos de captura de datos, ya sea porque se realicen de manera manual o si se capturan en sistemas de información, porque son finamente personas las que realizan estas actividades y por esta razón está sujeta a cualquier tipo de error. Resalta que indudablemente al hacer la captura manual, el número de errores es mayor por la falta de controles. Por su parte, el experto entrevistado 2 atribuye a las causas, el desconocimiento de los conceptos de calidad de datos por parte tanto de las organizaciones como de las áreas que realizan este tipo de labores enfatiza que el desconocer cuál es el valor de los datos y cuál es el estado de estos, los lleva a convivir con este problema por mucho tiempo sin tener mayor entendimiento del por qué sucede esto. El experto 1 entrevistado por su parte, atribuye a las causas, la falta de controles como reglas de depuración, homologaciones, estandarización y en general a la ausencia de políticas reales que mitiguen las implicaciones de los procesos en gestión de datos.

Por último al contrastar los dos instrumentos utilizados para este diagnóstico, se comprueba algunas de las afirmaciones que, desde la perspectiva profesional del autor del presente trabajo de grado planteó en el apartado de la problemática, encontrando inclusive que las estimaciones de tiempos empleados por los responsables para la gestión de los datos son mayores a las que se percibían, como lo expresa uno de los responsable de gestión entrevistado *“me puedo gastar toda la semana, solamente con TI duré casi 3 semanas tratando de ajustar la información para poderla enviar en lo posible.”* esto por los reprocesos manuales para el ajuste de las bases.

Adicionalmente, con la evaluación de la matriz se comprueba que las faltas más comunes que se han venido detectando sobre los datos, son: oportunidad, validez, integridad y exactitud.

De acuerdo con el anterior diagnóstico se pueden identificar las siguientes causas que generan la problemática actual: Ausencia de lineamientos para los operadores de servicio en los procesos de captura de datos, falta de lineamientos en la entrega de los datos que reportan, falta de controles de calidad en la revisión de los datos recibidos, falta de procedimientos establecidos para la gestión de calidad de los datos, falta de alineación con las demás áreas consumidoras de la información y por último falta de una cultura organizacional entorno a la gestión de datos que permita concientizar y generar capacidades en las personas responsables de este tipo de procesos en la Entidad.

4.2 Criterios relevantes y buenas prácticas de los marcos de referencia a aplicar

En la presente sección se va a evaluar y seleccionar criterios relevantes y buenas prácticas de los marcos de referencia mencionados en el presente trabajo, para adaptarlos en un plan que atienda las necesidades de gestión de datos de la oferta institucional en la Dirección de Apropiación.

Para la identificación y selección de los criterios relevantes de los marcos referencia señalados en el presente trabajo de grado, se aplicaron las guías de entrevista a los profesionales expertos en las cuales se les preguntó con base en su experiencia cuales criterios de estos marcos consideraban relevantes para la problemática planteada.

El experto entrevistado 1, hace referencia a la necesidad de definir una política que permita establecer las exigencias mínimas de calidad que deben presentar los datos, definición de métricas para la medición y control continuo ya que resalta en sus palabras “las decisiones que se lleguen a tomar usando información no confiable, representan un sesgo y un riesgo para las compañías, “...si tu calidad de datos es del 40%, tienes el 60% de margen de cometer errores en la toma de decisiones”. En cuanto al marco de referencia que por experiencia y por ser partner de Microsoft, sugiere la utilización del *Enterprise Information Management*, es un marco poco conocido, pero que en su aplicación le ha dado buenos resultados. Esta afirmación “...no hay nada más peligroso que una persona sin capacitación, sin las cualidades mínimas requeridas... tanto en el proceso, como en datos” deja ver que éste experto, encuentra necesario que, en el marco de la gestión de datos, los responsables se capaciten, se alineen con los objetivos

estratégicos del área y la organización a la que pertenecen de tal manera que comprendan a cuáles procesos misionales está apoyando con las actividades que realiza.

El experto entrevistado 2, resalta la importancia de que, en la implementación de planes de calidad, lo primero que se debe realizar es la identificación de las fuentes de información y lo segundo es la priorización de estas, como tercer punto que lista es tener un modelo de gobierno de datos para saber “...cuándo, cómo y dónde extraigo la información y la gestiono”; dentro de ese gobierno de datos, sugiere implementar lo que él denomina la “práctica de calidad de datos” que es todo el ciclo de extracción, perfilamiento, aplicación de reglas de negocio y mejoramiento o enriquecimiento de las fuentes de datos.

En cuanto al criterio para seleccionar un marco de referencia adecuado, señala que todos los marcos permiten llevar a la organización a una transformación digital, dice que los marcos ni son buenos ni malos, por lo que considera que el marco más recomendable a implementar, es aquel que más se ajuste a las necesidades de la organización y el que más impacte sobre la calidad de los datos, o que bien pudiera ser la mezcla de varios, simplemente no se casa con ninguno.

En cuanto a las personas, este entrevistado clarifica que es necesario que los directamente involucrados en estos procesos tengan **actitud**, más que **aptitud** para desempeñar su labor, sumando a que, si tienen bien definido y delimitadas sus responsabilidades, podrán desempeñar de manera adecuada sus funciones, por lo que encierra al final “...actitud, conocimiento, y responsabilidades definidas” como factores clave. Éste experto en especial, quien trabaja como parther para una de las consultoras más importantes a nivel Internacional, durante la mayor parte de la entrevista, enfatiza mucho en la importancia de la implementación de un marco de gobierno para la implementación o adopción de cualquier proyecto relacionado con gestión de datos, una de sus frases más destacada encerrando todo este concepto es: “Los datos, si no tengo procesos y si no tengo una metodología al final, no lo puedo gobernar y lo que no se controla, lo que no se gobierna, no tiene mejora continua”; esto lo resalta definiendo el valor del dato como monetizable mediante un modelo de gobierno de datos, y que las organizaciones tanto privadas como públicas pueden llegar a obtener este valor si saben en dónde están sus datos, si saben cómo están esos datos y para qué les sirve. Solo así podrán tomar decisiones basadas en el valor del dato para cumplir con los objetivos estratégicos, pero esto se logra a través de la adopción de un marco de gobierno de datos.

El experto entrevistado 3 considera como criterio fundamental específicamente para la Dirección, la estandarización y especificación de tipos de datos en los formatos de cara al operador concluyendo con sus palabras “...Entre más específico sea el dato que se capture, más específico sea al formulario para el

que se lo solicita pues menor la posibilidad de error y segundo, las personas que capturan deberían obedecer a procesos un poco más estandarizados, por eso hay procesos de calidad". Frente a los marcos, menciona que si bien por el rol que ejerce se encuentra más enfocado a procesos de analítica e inteligencia artificial, estas soluciones no se llevan a cabo con éxito si la data no presenta calidad, por lo que sí recomienda usar marcos que sean más específicos en el tema de gestión de datos y no las ISO, ya que considera que éstas se abordan de manera mucho más generalizada, y que eventualmente luego de que las organizaciones hayan adoptado marcos de gestión de datos, pueden evolucionar en cumplimiento de una ISO.

El experto entrevistado 4 menciona que dentro los criterios relevantes para tener en cuenta en la adopción de un plan de gestión, es clave la estandarización del proceso de captura, es decir debe ser claro y debe estar definido previo a la actividad de captura de datos, definir cuales se van a capturar, definir los tipos de datos esperados para esos campos, identificar cuáles son los realmente obligatorios para tener una data completa y con valor, también resalta la importancia de que las personas encargadas de realizar la labor de gestión de datos, tengan conocimientos del negocio con relación al proceso que realizan, y que entiendan la importancia y su nivel de madurez.

En cuanto a marcos de referencia, recomienda que la guía del DAMA es el más popular y el más conocido, siendo que es una guía de referencia con una serie de pasos que le permite a cualquier organización realizar una gestión adecuada de sus datos. El experto 4 menciona que a lo largo de su experiencia en el ámbito de TI, existe un comportamiento generalizado en las organizaciones en las que se presenta que, las iniciativas de los proyectos que tengan que ver con tecnología, son promovidos desde el área de TI; sin embargo, por su experticia considera que es un error que se presenta por la conceptualización que se tiene de la tecnología, y que al final del ejercicio este tipo de proyectos fracasa por la falta de alineación con la estrategia y por falta de dolientes y patrocinadores, por lo que su recomendación puntualmente en temas de gestión de datos, es que no debe ser una iniciativa únicamente de TI, sino que debe ser una labor conjunta ya que desde TI se entiende la importancia pero se debe comunicar e involucrar a las áreas de negocio para que participen activamente y para que los proyectos puedan ser exitosos.

En general todos los expertos coinciden en que el mejor marco para adoptar es aquel que se ajusta a las necesidades puntuales de cada de empresa y que le permita avanzar en la transformación digital, puede ser una mezcla de varios referentes para que se adapte y adopte a las necesidades organizacionales, dado que éstas varían de acuerdo con el tipo y cultura de la entidad a implementar. Se puede destacar que 3 de los 4 expertos entrevistados, enfocaron sus respuestas en

cada uno de los bloques temáticos siempre pensando en las necesidades del negocio, sus objetivos estratégicos y como apalancarlos a través de este tipo de proyectos. El otro experto en datos se analiza más de manera técnica, no queriendo decir con esto que no hubiese aportado a la investigación, pero su aporte por las respuestas entregadas presentaba un nivel mucho más enfocado a lo técnico que de igual manera fue provechoso para el ejercicio.

A continuación, se relacionan los criterios seleccionados para tener en cuenta en el plan, luego de evaluar su soporte teórico en el marco conceptual:

- Definición de un marco de gobierno de datos.
- Identificar los datos que son más importantes para la Entidad.
- Estandarizar formatos de captura.
- Identificar los procesos del ciclo de vida del dato.
- Estandarizar de procesos.
- Estandarizar procedimientos.
- Definir actividades, funciones, roles y responsabilidades.
- Capacitar a las personas frente a la importancia misional de sus funciones.
- Habilitar las capacidades técnicas y conceptuales de la labor que realizan.

Dado lo anterior y mediante la triangulación realizada a partir de las conclusiones de los bloques temáticos (Entorno ideal para la gestión de los datos) y (Marcos de referencia) y el marco conceptual se infiere que, podemos evidenciar con total claridad que lo presentado el marco conceptual del presente trabajo de grado, es coherente con los criterios y buenas prácticas que los expertos entrevistados recomiendan como clave a tener en cuenta para adaptar y adoptar en una estrategia de calidad de datos para la Entidad.

4.3 Acciones de mejora propuestas a los procedimientos actuales

En esta sección se proponen las acciones de mejora a los procedimientos actuales de gestión de datos de la Dirección de Apropiación, haciendo la claridad que no existen procedimientos formales sino tácitos.

Durante el proceso de análisis de las entrevistas a los responsables de la gestión de datos en cuanto a los procedimientos actuales que realizan, se logra identificar que las exigencias a los operadores para la entrega de información al MinTIC se basan únicamente en la definición de campos que consideran que se requieren desde el punto de vista de negocio, desconociendo las demás consideraciones técnicas de vital importancia para garantizar la calidad de los datos.

Dentro de los procesos que realizan para la verificación de las bases entregadas, todos manifiestan realizarlos de manera manual, reconociendo una parte de los problemas en los datos de manera empírica o "...a ojo" como lo describe uno de los entrevistados, la detección de inconsistencias, la revisión de la validez de la información es realizada sin ningún tipo de documentación o herramientas que les ayuden a ejecutar estas actividades. De acuerdo a lo narrado por uno de los responsables, comenta que en un momento por iniciativa propia planteó en una matriz con unos campos estándares para la identificación de la población que se está beneficiando, lo que comúnmente se denomina como enfoque diferencial o caracterización de la población, dado que estos requerimientos suelen realizarse a las áreas por derechos de petición o para informes que presenta el gobierno, sin embargo, esta iniciativa no tuvo éxito, ya que no fue apalancada desde la estrategia, por lo que las demás áreas desconocen la importancia de este tipo de herramientas que aportan a la solución de la calidad de datos en la Entidad.

De manera repetitiva, los inconvenientes que encuentran estos gestores de información, son solucionados como pueden, duplicando esfuerzo para dar cumplimiento a las demás áreas que requieren de la información para la toma de decisiones, en muchas ocasiones, acuden a los mismos operadores de servicio a fin de corregir inconsistencias encontradas, sin embargo, esto se convierte en un inconveniente adicional por los temas contractuales acordados al inicio, lo que no obliga al operador a ajustar la información entregada, de acuerdo con cada cambio que el Ministerio pueda requerir. Por último y con base en la información proporcionada por los entrevistados, todos ellos manifiestan no conocer algún plan, procedimiento, manual o guía formal dentro del Ministerio del que puedan hacer uso para las actividades que realizan a diario, afirmación que se comprueba mediante triangulación realizada entre la revisión del MIG (Modelo Integrado de Gestión) y el bloque temático 2 (*Ver anexo 7*).

De acuerdo con lo anterior, se concluye que la propuesta está basada en las siguientes acciones de mejora a los procedimientos actuales:

- Definir un procedimiento de Calidad de Datos en la Entidad que sea, medible, auditable y susceptible de mejora.
- Definir estándares o lineamientos para la captura de datos tanto a nivel de negocio como a nivel técnico, en las etapas contractuales con los operadores de servicio.
- Apropiar alguna herramienta tecnológica que facilite la detección de errores asociados a la calidad de los datos.
- Apropiar de alguna herramienta tecnológica que facilite el ajuste a los datos que sean detectados en la etapa de revisión.
- Estandarizar de reportes de información a los diferentes interesados.

Estas propuestas de mejora estarán inmersas en la propuesta del plan en el capítulo 5 del presente trabajo de grado.

4.4 Identificación de los beneficios de un plan de gestión de calidad de datos

En la presente sección se identifican los beneficios de proponer un plan de gestión de calidad de datos para la Dirección de Apropiación.

“...Porque algo peor que tener mala calidad en los datos, es no saber que se tiene mala calidad de datos.”, así lo considera uno de los expertos entrevistados. Por esa razón, destacan la importancia de que exista en las organizaciones alguna estrategia que les permita identificar y disminuir las fallas que presentan los datos, a fin de tener una mejora continua para seguir el curso hacia la transformación digital apalancada en la toma de decisiones acertada.

Todos los expertos coinciden en que la adopción de un plan de calidad de datos se considera uno de los aspectos básicos a tener en cuenta en cualquier organización dentro de su arquitectura de datos, de manera que a través de su adopción le entregue valor al dato para ser usado de manera asertiva acorde con la misionalidad de la Entidad. Esto con el fin de lograr que el dato sea monetizable, lo que no hace referencia al dinero, sino a que obtenga mayor valor en la medida en que estos puedan ser utilizados para la consecución de los objetivos de la organización, por lo tanto, para lograrlo se requieren de proyectos e iniciativas que se promuevan desde un plan establecido y que con el tiempo abra paso a uno de los beneficios más importantes que se pueden obtener, el cual consiste en generar una madurez cultural en la organización en la medida que este sea apropiado por todos involucrados en los diferentes procesos de la gestión de tal modo que de manera casi que autónoma los responsables de la gestión, velen por la calidad de sus datos entendiendo la importancia de estos, lo que produce a mediano plazo una mejora en la eficiencia de las actividades diarias volviéndolas más realizables, fáciles y a su vez optimizando los reportes de información que son proporcionados a la estrategia para la toma de decisiones.

Por último y desde la perspectiva de los responsables de gestión de datos que son quienes se encuentran sumergidos en esta problemática, consideran que la adopción de un plan de gestión para la calidad de datos les facilitaría los procesos que realizan actualmente, evitando los reprocesos manuales, impactando el recurso financiero mediante la disminución de costos y a su vez de tiempo, que pueden ser usadas en actividades más enfocadas al cumplimiento de los objetivos estratégicos.

De acuerdo con estas posiciones (expertos y responsables de gestión de datos), se podrían resumir los beneficios de la siguiente manera:

- Generación de datos más acertados.
- Agilidad en la gestión.
- Disminución de tiempos.
- Disminución de costos.
- Reducción de reprocesos.
- Cambio cultural.
- Madurez institucional.
- Valor agregado en los datos.
- Toma de decisiones asertivas.

5 Diseño y propuesta de implementación del plan de gestión de calidad

El presente capítulo aborda el diseño y propuesta de implementación del plan de gestión de calidad de datos, basado en las acciones de mejora propuestas y los criterios seleccionados de los marcos de referencia de gestión de datos.

Para la propuesta del plan de gestión de calidad, no solo se tendrán en cuenta los criterios seleccionados de los marcos de referencia y las acciones de mejora a los procesos actuales, sino que con base en el análisis de las entrevistas realizadas a los expertos se tienen en cuenta las siguientes:

- Conocimiento del negocio (Estrategia, Misión y objetivos de la organización)
- Definición de un modelo de gobierno de datos.
- Identificación de las fuentes de información y su importancia para la organización.
- Diagnóstico de calidad de datos.
- Roles y responsabilidades.
- Procedimientos.
- Capacidades técnicas y humanas.
- Indicadores de Medición.

Acorde al estado actual de la Dirección de Apropiación, se analizaron los procesos actuales y se alinearon con los criterios de referencia seleccionados con el propósito de adaptarlos a la solución planteada para mejorar la gestión de datos.

5.1 Objetivo

Mejorar la oportunidad y pertinencia de la información de la oferta institucional en la Dirección de Apropiación del Ministerio TIC de tal forma que soporte la toma de decisiones estratégicas en la Dirección y de la estrategia que gira en torno a esta área.

5.2 Alcance

El plan incluye la definición de un modelo de Gobierno de Datos basado en las recomendaciones de (Gati Consultores) como uno de los puntos más relevantes para la gestión de datos en la Dirección de Apropiación y la identificación de los procesos involucrados, cuyas fuentes son provistas de forma manual en archivos

digitales por los operadores de servicio de los diferentes proyectos en ejecución. Dichos puntos incluyen: identificación de procesos, flujo de información (ciclo de vida del dato), procedimientos de calidad, definición de roles y responsabilidades, forma de medición y herramientas de apoyo clave para incrementar y mantener los niveles de calidad esperados que soporten los procesos del MIG (Modelo Integrado de Gestión) de la Dirección de Apropiación.

5.3 Estrategia de gobierno de datos

Los datos son el activo más importante para la Dirección de Apropiación en la toma de decisiones para alcanzar el cumplimiento de los objetivos estratégicos y tácticos, por esta razón la estrategia de implementación del modelo de gobierno de datos que se recomienda adoptar está enmarcada en las siguientes actividades macro que debe realizar la Entidad una vez decida adoptar el presente plan:

1. **Priorización de los datos críticos:** es una actividad fundamental que ayuda a decidir qué es más urgente, más importante, cuáles son los datos críticos para el área acorde con la estrategia de la entidad y cuál es la ruta de trabajo a seguir para gobernarlos. Para ello, son clave las siguientes actividades:

- Identificar las necesidades o preguntas de negocio a resolver,
- Definir las entidades de negocio o dominio de datos a gobernar,
- Identificar los grupos de atributos,
- Identificar los procesos de negocio involucrados,
- Identificar los sistemas de información relacionados.

Los procesos de negocio son seleccionados con base a la matriz CRUD, donde se identifica qué tratamiento realizan sobre el dato, utilizando operaciones básicas como: creación, lectura, actualización y eliminación.

2. **Definición de roles y responsabilidades:** este paso consiste en identificar y definir los roles del proceso de gestión de datos, se realiza con el fin de establecer la estructura jerárquica de responsabilidades y nivel de autoridad para la toma de decisiones en asuntos relacionados con el estado de los datos a gobernar. Los roles principales y más comunes en gobierno de datos son: el líder de gobierno de datos, los administradores y propietarios de los datos, estos roles se definen en el marco de adopción del plan, y deberá ser el resultado de las reuniones que se realicen de manera conjunta con las áreas involucradas.

3. **Implementar la práctica de calidad de datos:** por último, en este paso, se pretende generar las capacidades para la implementación de la práctica de calidad de datos, con el fin de garantizar que estos son de alta calidad para su uso, promoviendo a través de cada uno de los responsables de la gestión de datos en la Dirección de Apropiación, la conciencia y por ende la cultura de calidad de datos. El ciclo de vida de esta práctica se encuentra definido en el numeral 5.5.2 en el proceso de calidad.

5.3.1 Impulsores de negocio

Los impulsores de negocio también conocidos como motivadores, son los conductores vitales para el éxito continuo y el crecimiento del negocio, dado que estos deben estar alineados con la visión estratégica de la entidad y de la Dirección de Apropiación, a fin de que puedan generar valor desde la implementación del marco de gobierno hacia la estrategia. A continuación, el autor del presente trabajo recomienda a MinTIC considerar en priorizar los siguientes impulsores:

Nota: los marcados con asterisco (*) son los que se sugiere priorizados.

- Los datos como un activo de negocio estratégico (*)
- Sobrecostos operativos ocultos
- Mejorar la toma de decisiones basada en datos (*)
- Cumplimiento de la ley y cambios en regulaciones.
- Generación de datos de manera eficiente y precisa para uso de cualquier ciudadano a través del portal de Datos abiertos del gobierno colombiano (*)

5.3.2 Misión

La misión propuesta se refiere a la definición de la razón de ser del gobierno de datos. Es aquello que le da foco y permite tomar decisiones de las acciones que debe hacer, optimizando el uso de los recursos. Se propone la siguiente definición para la misión del gobierno de datos:

Somos el órgano de gestión de datos en la entidad que administra, mejora y pone a disposición los datos con mayor calidad hacia los consumidores clave, mediante esfuerzos coordinados para mejorar la eficiencia y la toma de decisiones.

5.3.3 Visión

La visión permite señalar hacia dónde se dirige el gobierno de datos como función, marcando el horizonte de tiempo en el mediano y largo plazo e identificando las actividades que deben ser realizadas para alcanzarla. Se propone la siguiente definición para la visión del gobierno de datos:

En 2021 el gobierno de datos será reconocido por elevar el estado de los datos priorizados como un activo de negocio fácilmente disponible para respaldar la toma de decisiones a través de metodologías que permitan administrar los datos de la mejor manera en conjunto con los funcionarios de la Dirección de Apropiación.

5.4 Objetivos, principios y políticas de gobierno de datos

Como objetivo se denomina el fin al que se desea llegar o la meta que se pretende lograr implementado gobierno de datos en el negocio.

5.4.1 Objetivos de gobierno de datos

La Dirección de Apropiación debe tener en cuenta en su proceso de adopción del plan, los siguientes objetivos para el gobierno de datos que se implemente:

- Crear políticas eficientes y efectivas en torno a los datos de los proyectos estratégicos de la Dirección, según entrevista a expertos.
- Incrementar la consistencia y la confianza en la toma de decisiones basada en los datos de la Dirección.
- Aumentar el valor de los datos priorizados, en este caso los proyectos de la “Oferta Institucional” de la Dirección de Apropiación.
- Determinar y asignar las responsabilidades de gestión de datos en la Dirección de Apropiación.
- Apalancar el cumplimiento de los objetivos estratégicos de la Dirección de Apropiación.
- Mejorar la comunicación entre la Dirección de Apropiación y la oficina de TI gestionando su integración.

5.4.2 Principios de datos

A continuación, se relacionan los principios recomendados para promover el buen gobierno de datos, una vez la Entidad decida la adopción del plan, se sugiere que estos principios sean evaluados en conjunto con áreas de negocio, de TI y el oficial de seguridad de la Entidad, a fin de definir cuáles de estos se implementan de manera prioritaria; no obstante el autor del presente trabajo y desde su perspectiva por el conocimiento previo que tiene de la Dirección, sugiere tener en cuenta de primera mano los marcados en asterisco (*).

Principio No. 1 El dato es un activo y se gobierna (*)

- Los datos son propiedad del negocio y deben gestionarse como un activo clave para la entidad.

Principio No. 2 El dato está disponible, se accede y es seguro.

- Los datos se encuentran a disposición de quienes deben acceder a ellos, usuarios de negocio, procesos, aplicaciones de acuerdo con las necesidades y roles definidos.
- Se garantiza el uso adecuado de los datos por las partes interesadas y se limita el acceso a datos sensible o confidenciales acorde con la normatividad vigente.

Principio No. 3 Datos con calidad (*)

- Los datos son completos, correctos, reflejan la realidad y son aptos para su uso porque cumplen con un conjunto de características definidas como, por ejemplo: completitud, conformidad, consistencia, exactitud, duplicidad e integridad.
- Se debe evitar la duplicación innecesaria de los datos en las fuentes origen.

Principio No. 4 El dato es auditable (*)

- La auditoría de calidad a los datos se realiza con el fin de verificar cumplimiento de las reglas de negocio establecidas.

Principio No. 5 Decisiones basadas en Información (*)

- Todas las decisiones estratégicas y tácticas se soportan en datos de calidad, debidamente gestionados, auditados, controlados y publicados; para asegurar así la toma de decisiones de manera objetiva y coherente.

Principio No. 6 Resolución de problemas

- La resolución de problemas seguirá la estructura definida en el gobierno de datos para controlar los derechos de decisión y autoridad para asuntos relacionados con los datos.

Principio No. 7 Toma de conciencia (*)

- Todos los involucrados relevantes interiorizaran sobre la importancia del manejo de los datos, así mismo valoran el rol estratégico de la información; siendo conscientes de los datos y la información a su cargo, así como de los procesos de negocio que los crean, almacenan, procesan y utilizan.

5.4.3 Políticas de datos

De acuerdo con uno de los expertos entrevistados el cual hacía énfasis en la importancia de un gobierno de datos, y mencionaba acerca de conocer quién, cómo y cuándo realiza qué con los datos. Esa expresión traducida en términos de un marco de referencia como el DAMA, hace referencia a las políticas; las cuales son las que describen el “qué” del gobierno de datos, es decir, qué hacer y qué no hacer.

Las políticas que se recomiendan en este plan se derivan de la estrategia de gobierno de datos que se sugiere implementar, estas políticas deben comunicarse, supervisarse, aplicarse y actualizarse periódicamente, se recomienda mínimo (1) una vez al año acorde con los procesos de auditoría internas que se realizan en la Entidad.

Se debe tener en cuenta que las políticas de datos toman diferentes formas acordes con los dominios de datos priorizados a los que éstas van enfocadas. Si la Dirección de apropiación decide en el marco de la adopción del plan priorizar los datos de la Oferta Institucional una política pudiera ser:

- Para todos los datos del dominio de negocio “Oferta Institucional” certificados, se debe evaluar cada 6 meses su precisión, integridad, coherencia, accesibilidad, cumplimiento y eficiencia.

A continuación, se describe la **estructura propuesta de una política** de gestión de datos, la cual se sugiere sea tomada en cuenta para la definición de estas:

Política de retención y eliminación de datos:

Propósito

El propósito de esta política es garantizar que los datos utilizados por la Dirección de Apropiación se archiven o se eliminen de manera coherente, al tiempo que se garantiza que no se violen las leyes, la seguridad o la privacidad en la disposición.

Declaración

- Los propietarios de datos identificarán claramente los períodos de retención para cuando los datos deben estar disponibles para el acceso diario dentro de los sistemas de la entidad, así como para fines de archivo de acuerdo con la ley y las necesidades internas.
- El período de retención puede variar según las leyes específicas aplicables en nuestro país.
- Al alcanzar el período de retención, todos los datos deben archivarse primero durante un período específico según lo determinen los propietarios de datos del dominio de datos.

Aplicación y seguimiento

La presente política aplica para los datos del dominio de “Oferta Institucional”.

Tabla 18 Estructura propuesta de una Política de Gestión de Datos

Fuente: Elaboración propia

Si bien las políticas planteadas son apenas una propuesta, el equipo de Gobierno junto con la Dirección de Apropiación, debe definirlas y refinarlas. A continuación, se presenta la propuesta de las políticas que deben ser desplegadas en la Entidad para consolidar el modelo propuesto:

Política de acceso a datos: el propósito de la política de acceso a los datos es garantizar que todas las áreas de negocio que interactúan con los datos de la Dirección de Apropiación tengan el nivel adecuado de acceso a los datos según las

necesidades tácticas y operativas requeridas para el desarrollo de las funciones asignadas.

Política de uso y divulgación de datos: el propósito de la política de uso y divulgación de datos es evitar que los datos de la Dirección de Apropiación se utilicen indebidamente para fines ilegales.

Política de privacidad y seguridad: dado que los programas de la Dirección de Apropiación son enfocados directamente a personas, la captura de sus datos debe estar debidamente protegida.

El propósito de la política de privacidad y seguridad de los datos es garantizar que los datos de la Dirección de Apropiación estén protegidos contra amenazas internas y externas y que cumplan plenamente con los requisitos de la ley 1581 de 2012 protección de datos personales mencionada en el marco de normativo del presente trabajo. Esta política debe abordarse junto con las políticas de acceso, uso y divulgación de datos.

Política de integridad (calidad de datos) e integración de datos: el propósito de esta política es garantizar que los datos utilizados en la Dirección de Apropiación tengan un alto grado de integridad y que los elementos de datos clave puedan integrarse con las demás áreas de negocio y otras aplicaciones para que los funcionarios puedan confiar en los datos para toma de decisiones.

La integridad de los datos se refiere a la validez, confiabilidad y precisión de los datos. La integridad de los datos se basa en una comprensión clara de los procesos de negocio subyacentes a los datos y la definición coherente de cada elemento de datos.

La integración de datos, o la capacidad para combinar datos de diferentes fuentes de información para convertirlos en información útil y valiosa.

Nota: estos objetivos, principios y políticas deben especificarse de manera detallada acorde con la estructura propuesta, una vez la Entidad entre en proceso de adopción del plan; lo anterior, porque hay asuntos específicos que se deben definir de manera interdisciplinaria entre todos los involucrados a fin de que sean partícipes de este proceso.

5.5 Modelo operacional del plan de Gestión de Calidad de Datos

Después de tener definido un modelo de gobierno de datos y de identificar la calidad de los datos como parte de una de sus políticas fundamentales, la propuesta se enfoca en los aspectos operacionales de la calidad de datos que aborda el plan para la Dirección de Apropiación de la siguiente manera:

El modelo de gobierno propuesto es el punto de partida para mejorar la gestión de datos, más que establecer roles y responsabilidades provee la metodología que articula personas, procesos y tecnología con un enfoque basado en procesos que permitirá gestionar los datos como un activo de negocio.

A continuación, se ilustra el modelo de gobierno de datos a alto nivel en el que se observa gráficamente en donde se encuentra ubicada la práctica de calidad de datos:

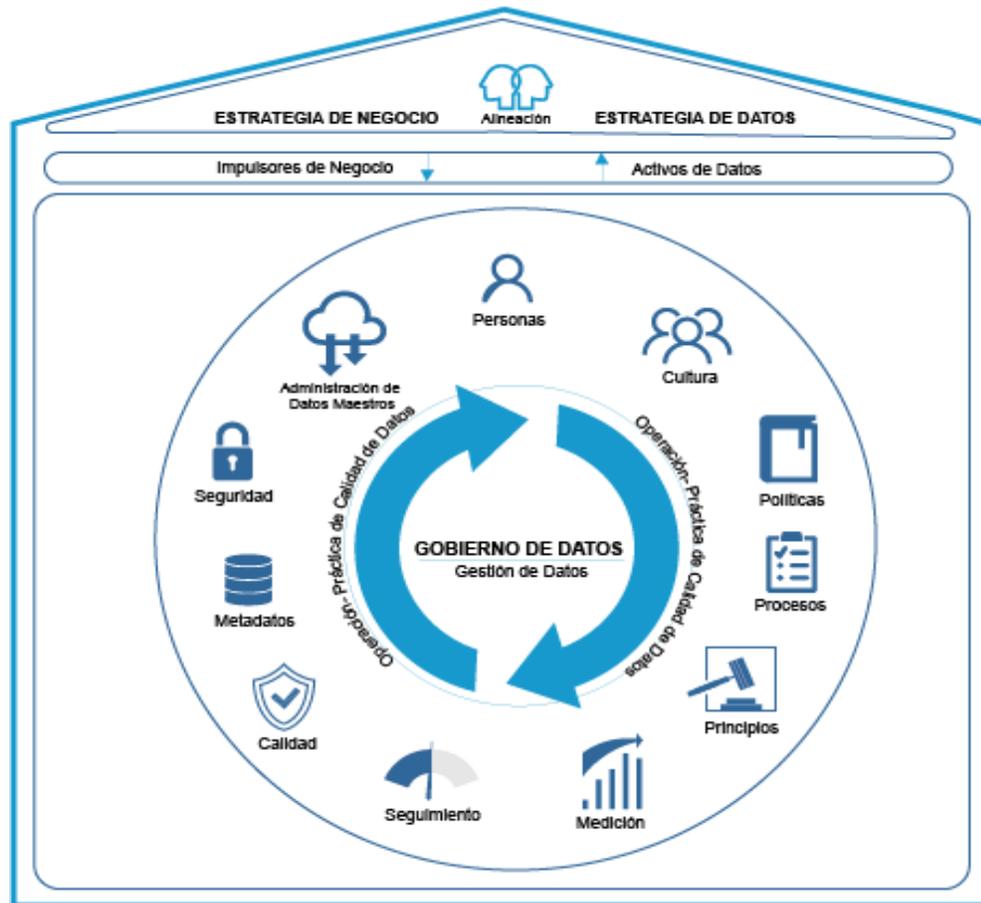


Figura 22 Modelo de gobierno de datos a alto nivel
Fuente: Elaboración Propia

Para operacionalizar el modelo se requiere desarrollar políticas, principios y procesos de gestión de datos los cuales describen las tareas y el paso a paso necesarios para planear, desarrollar, operar y controlar la gestión de datos.

El modelo es una herramienta de gestión que permite dirigir y evaluar el desempeño de gobierno de datos y se plantea con (2) capas así:

- **Capa de alineación:** la implementación del modelo alinea la estrategia de negocio y la estrategia de datos ya que habilita los datos de alta calidad necesarios para apalancar el cumplimiento de los objetivos estratégicos.
- **Capa de operacionalización:** define los procesos de necesarios para planear, desarrollar, operar y controlar la gestión de datos basado en las buenas prácticas que entrega el DMBOK. Este enfoque basado en procesos identifica y gestiona de manera eficaz numerosas actividades relacionadas entre sí.

5.5.1 Identificación del macroproceso de la Dirección de Apropiación

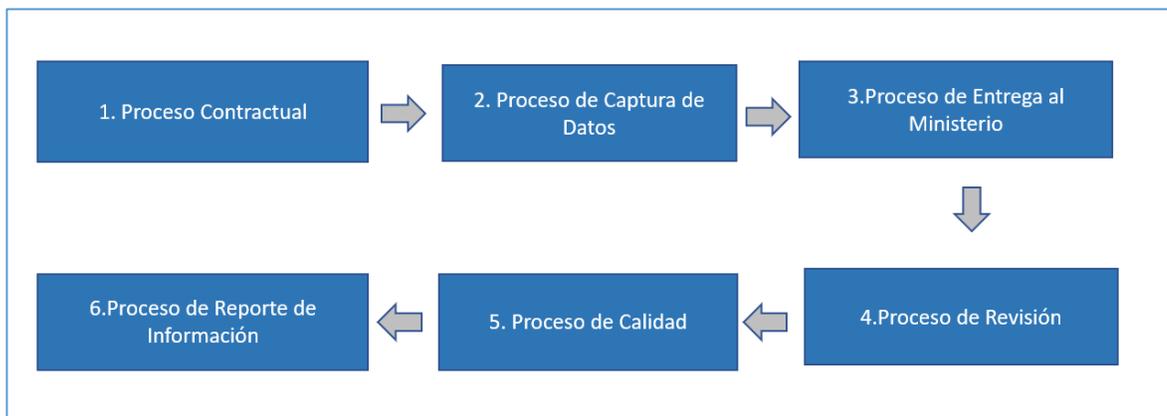
Dentro de la capa de alineación, se identifica que, la Dirección de Apropiación se encuentra dentro del Macroproceso Misional de la Entidad de *Acceso, Uso y Apropiación de las TIC*. Para efectos de la adopción del plan, éste debe ser implementado en el proceso de *Uso y Apropiación de las TIC*, señalado en la siguiente figura que busca dar cumplimiento de los objetivos estratégicos del Ministerio TIC en este ámbito.



Figura 23 Mapa de Macroprocesos del MinTIC
Fuente: Elaboración Propia

5.5.2 Procesos de gestión de datos en la Dirección de Apropiación

Dentro de la capa de operacionalización y en el marco de la investigación se identificaron 6 procesos clave en la gestión de datos sobre los cuales se debe desarrollar, operar y controlar la calidad de los datos mediante la aplicación del plan propuesto. Esta sección muestra por medio de un diagrama, los procesos formales (en el MIG) y tácitos involucrados y la manera como se deben abordar:



*Figura 24 Diagrama de flujos de los procesos formales y tácitos de Gestión de Datos en la Dirección de Apropiación
Fuente: Elaboración propia*

1 Proceso contractual: este es un proceso formal con el que cuenta el Ministerio en el cual se recomienda incluir un lineamiento a nivel de la entidad, en el que se promueva y estandaricen los procesos de recolección de datos para cada uno de los diferentes proyectos que la Dirección de Apropiación tenga programados ejecutar, acorde con el plan de acción del Ministerio TIC. Este lineamiento debe contemplar los requisitos de información que requiera el área de negocio, la forma de capturar los datos, las dimensiones de calidad de datos a tener en cuenta y las herramientas a utilizar para ser solicitadas al operador del servicio.

Se recomienda que la creación de este lineamiento sea impulsado por el Director(a) de la Dirección de Apropiación y definidos de manera interdisciplinaria entre áreas como: la Dirección de Apropiación (área de negocio), la Oficina de Planeación (área estratégica), la coordinación de contratación (área jurídica) y la oficina de tecnologías de información (área Técnica) en el marco de gobierno de datos propuesto.

Lo anterior, con el fin de que estos requerimientos de información hacia el operador a contratar sean concebidos desde la estructuración de los proyectos, y que desde cada uno de los involucrados se proyecte el impacto y beneficios de tener en cuenta de manera preventiva todas las necesidades que puedan surgir a raíz de la ejecución de un proyecto misional para cumplir con los objetivos de negocio

2 Proceso captura de datos: actualmente este proceso lo realizan los operadores de servicio basado en su propio criterio, por lo que se recomienda que

se derive del proceso número 1 (1 Proceso contractual) , por lo tanto, los operadores de servicio que contraten con la Dirección de Apropiación la ejecución de algunos de sus proyectos, deben cumplir a cabalidad con los lineamientos establecidos en el gobierno de datos y deben quedar plasmados en el contrato para la entrega de los datos capturados en dichos proyectos. De tal modo que las herramientas y estándares que utilicen en este proceso deben estar acorde con lo solicitado por el Ministerio y aceptado por el operador de servicio en contrato realizado, a fin de mejorar los resultados de los datos que entregan a la Dirección de Apropiación.

3 Proceso de entrega al Ministerio: este es un proceso tácito que realizan los operadores de servicio, por lo que se recomienda especificar de manera clara y concreta las condiciones en el proceso número 1 (1 Proceso contractual); de manera tal, que para la Dirección de Apropiación debe ser clara la forma, el medio y los tiempos de recepción de las bases de datos que contienen la información acordada.

4 Proceso de revisión: en este punto, los gestores responsables de los datos en la Dirección de Apropiación realizan a la fecha un conjunto de actividades empíricas, por ésta razón, se recomienda realizar una validación objetiva acerca de los datos que reciben por parte de los operadores, de acuerdo con las especificaciones de negocio, tipo de datos (dimensiones), tiempos y medios de entrega que se definan en los procesos previos; convirtiéndose en un proceso estandarizado que puede realizar cualquier funcionario o servidor público siempre y cuando, se base en los lineamientos definidos en el proceso número 1(1proceso contractual) en el marco de gobierno de datos.

Por otra parte, la revisión específica en cuanto a la conformidad de los datos recibidos debe realizarse con apoyo de herramientas tecnológicas que faciliten la ejecución de esta actividad; en el mercado existen numerosas herramientas que pueden soportar el apoyo de estas actividades y que son de uso libre, como también las hay licenciadas. No obstante, en el punto 5.5.7 del presente plan, se realiza la recomendación de una de estas herramientas.

5 Proceso de calidad: éste es un proceso que no se encuentra documentado ni formalizado en la entidad, por lo tanto y de acuerdo con los expertos entrevistados y a las recomendaciones de los marcos de referencia, se recomienda implementarlo en el modelo operacional de calidad de datos para poder realizar el seguimiento al ciclo de calidad de datos. El ciclo contempla los siguientes pasos:

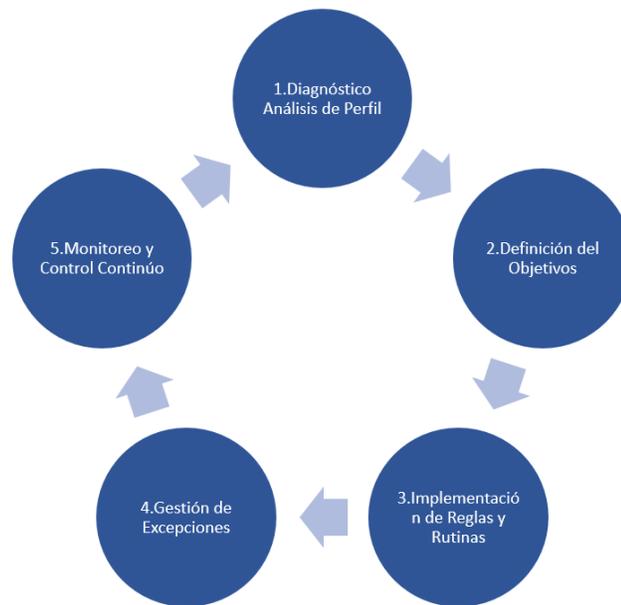


Figura 25 Ciclo del proceso de Calidad de los Datos
Fuente: Elaboración propia

- Diagnóstico y Análisis del perfil (Perfilamiento de datos): este es el punto que recomiendan marcos como el DAMA, para conocer patrones de comportamiento de los datos tales como nulos, vacíos, longitud de campos y tipos de datos; de igual forma es importante realizar el diagnóstico del conjunto de datos a evaluar, para conocer su estado, acciones de mejora a implementar y porcentaje de la calidad de datos actual.
- Definición de objetivos: en este punto, las áreas de negocio que para el caso es la Dirección de Apropiación, debe definir los objetivos que serán medibles a lo largo de este ciclo; es decir, el aumento de la calidad de los datos en un porcentaje determinado; no obstante, si los datos ya presentan un nivel óptimo de calidad, el objetivo debe estar orientado a mantener dicho porcentaje. En cualquiera de los dos casos, estos objetivos son establecidos acorde con las necesidades del negocio y con los campos definidos dentro del modelo de gobierno de datos como críticos para la Dirección de Apropiación.
- Implementación de Reglas y Rutinas: en este paso, se definen las reglas de negocio a validar sobre los datos que van a ser evaluados, los cuales ya han sido establecidos en la política de gobierno de datos. Acorde con esas reglas se definen las rutinas que deben ser ejecutadas para el cumplimiento de los objetivos, ya sea para adaptarlos de manera preventiva en la próxima iteración o correctivos para permitir su curso en el flujo de información.

- Gestión de Excepciones: si bien en los dos pasos anteriores, se definen tanto los objetivos como las reglas a implementar sobre los datos, la definición de excepciones es un punto clave, a fin de no entorpecer el flujo de los datos y su objetivo, por lo tanto, estas definiciones deben ser objeto de revisión por parte del área de negocio y un área técnica.
- Monitoreo y control continuo: este punto se realiza con base en la definición de los objetivos a medir en el ciclo de la calidad de datos, con la finalidad de poder realizar seguimiento a estos indicadores y conocer si se están cumpliendo, si han desmejorado o se han mantenido, para encontrar hallazgos y poder generar acciones de mejora. Es necesario hacer visibles estos indicadores para poder realizarles de manera más ágil el seguimiento.

6 Proceso de reporte de información: este proceso no es formal ni está definido con las demás áreas de la entidad, por lo cual se recomienda que la Dirección de Apropiación establezca acuerdos con las áreas consumidoras de los reportes; para ello, debe socializar los diferentes proyectos a ejecutar y detalles tales como objetivos, duración, población objetivo y demás información que permita previamente definir entre las partes el estándar de reporte acorde con los principios de los datos definidos en el marco del gobierno. Esto le permitirá a la Dirección realizar este proceso de manera estandarizada y más eficiente.

Para lograr esto, se propone en el numeral 5.5.4 Procedimientos propuestos para la Dirección de Apropiación, un procedimiento de reporte de información.

5.5.3 Flujo de la información en la Dirección de Apropiación

El siguiente es el flujo de información que se logró identificar producto de la investigación realizada, en el que se muestran puntos clave mencionados por los responsables de la gestión de datos en la Dirección de Apropiación y sobre los cuales se deben aplicar los controles establecidos en los procesos mencionados en el punto 5.5.2 del presente plan.

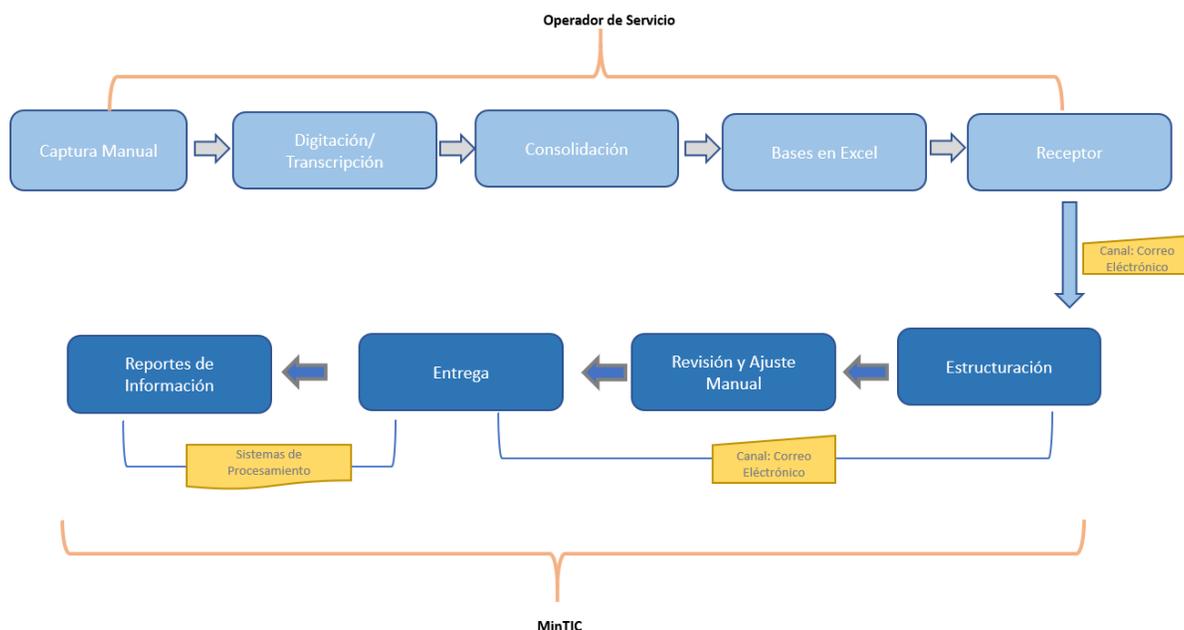


Figura 26 Flujo de la información en la Dirección de Apropriación
Fuente: Elaboración Propia

5.5.4 Procedimientos propuestos para la Dirección de Apropriación

Los procedimientos indican la forma en que se debe llevar a cabo todas las actividades que conforman los procesos, es en donde se muestra el paso a paso para poder ejecutarlas. El uso de procedimientos genera un mayor dominio de parte del usuario con respecto a estas actividades, evitando o mitigando posibles errores.

Dado el alcance del presente trabajo de grado, de los procesos mencionados en la sección 5.5.2 anterior, se van a proponer **procedimientos** únicamente para el **proceso de calidad y el proceso de reporte de información**, estos contemplan el enfoque coordinado para definir, medir y mejorar la calidad de los datos, y la manera como debe articularse el reporte de información, asegurando su idoneidad para el propósito y la alineación con las necesidades de la Dirección de Apropriación.

Por lo anterior, se recomienda establecer procedimientos de gestión de Calidad de datos y reporte de información en la entidad, que sean susceptibles de auditar y/o medir, los cuales deben estar basados en el ciclo de calidad de datos mencionado en el proceso número 5 *Proceso de Calidad* del punto 5.5.2. y en el 6 *Proceso de reporte de información*; para ello, se sugieren los **siguientes procedimientos**:

- **Diagnóstico de calidad de datos:** dar a conocer las actividades para evaluar, medir y monitorear los niveles de calidad de las fuentes de información, garantizando una gestión proactiva de la calidad de los datos e incorporando buenas prácticas que permitan precisar si los datos son aptos o no para el objetivo de negocio planteado.
- **Reparación de datos:** dar a conocer las actividades para corregir los errores identificados en el diagnóstico, a través del diseño e implementación del proceso de mejora de la calidad de los datos que comprende: limpieza, estandarización y enriquecimiento.
- **Enriquecimiento de datos:** dar a conocer el paso a paso para completar y ampliar los datos de los campos de las fuentes de información que mantienen baja calidad, a través de la comparación o cruce de datos con fuentes de confianza o referencia.
- **Reporte de información:** dar a conocer las actividades principales que debe tener en cuenta la Dirección de Apropiación, para reportar la información a los diferentes usuarios consumidores.

La siguiente tabla, define las actividades y actores principales que se recomiendan tener en cuenta para la creación detallada del paso a paso de los procedimientos propuestos:

Procedimiento	Actividades Principales	Responsables	Entradas	Salidas
Diagnóstico de calidad de datos.	<p>Priorizar el dominio de datos.</p> <p>Seleccionar las dimensiones de calidad a evaluar.</p> <p>Definir objetivos.</p>	<p>Área de negocio.</p> <p>Área técnica en conjunto con área de negocio.</p> <p>Área de negocio.</p>	<p>Bases de datos de la Dirección de Apropiación.</p> <p>Recomendaciones técnicas respecto a las dimensiones a seleccionar.</p> <p>Dominio de datos y dimensiones de calidad a evaluar.</p>	<p>Campos seleccionados de las bases de datos.</p> <p>Listado de dimensiones de calidad a evaluar.</p> <p>Indicadores de medición de calidad de datos.</p>

Procedimiento	Actividades Principales	Responsables	Entradas	Salidas
Reparación de datos.	Definir reglas de negocio.	Área de negocio.	Dominio de datos evaluados en el diagnóstico.	Reglas de negocio para aplicar a los datos.
	Definir rutinas de ejecución.	Área de negocio en conjunto con área técnica.	Reglas de negocio definidas y dominio de datos.	Rutinas de reglas de negocio estandarizadas.
	Implementar reglas y rutinas.	Área técnica.	Reglas de negocio con las rutinas definidas para los dominios de datos.	Datos con calidad.
Enriquecimiento de datos.	Identificar datos de referencia.	Área técnica en conjunto con áreas de negocios de la Entidad.	Dominio de campos priorizados y diagnosticados.	Fuentes de información con datos similares.
	Identificar datos maestros.	Área técnica en conjunto con áreas de negocios de la Entidad.	Dominio de campos priorizados y diagnosticados.	Fuentes de información con datos maestros.
	Cruzar las fuentes de información.	Área técnica.	Fuentes de información de confianza o referencia.	Datos enriquecidos.
Reporte de Información.	Identificar los proyectos nuevos.	Área de negocio.	Programas e iniciativas propuestos de la Dirección de Apropiación.	Listado de proyectos a ejecutar.
	Socializar con los interesados los proyectos.	Área de negocio.	Objetivos, duración, población, indicadores.	Necesidades de información de las áreas consumidoras.
			Necesidades de información de las	Listado de campos a

Procedimiento	Actividades Principales	Responsables	Entradas	Salidas
	Identificar los datos a solicitar al operador.	Área de negocio – Áreas consumidoras.	áreas consumidoras e indicadores.	solicitar al operador para generar reportes.
	Diseñar los reportes.	Área de negocio.	Estructura de datos definida que captura el operador.	Diseño de reportes estándar.
	Implementar Reportes.	Área de negocio - Área de TI	Formatos estándar y diseños propuestos.	Reportes funcionales estandarizados.

Tabla 19 Tabla de procedimientos propuestos
Fuente: Elaboración Propia

5.5.5 Medición del plan de calidad en el marco de gobierno de datos

La medición de los objetivos propuestos se debe realizar con base en el listado de campos críticos que hayan sido seleccionados por el área de negocio vs las dimensiones a evaluar en el momento de la adopción del plan y deben estar definidos al inicio de la implementación del plan de calidad de datos.

Para esta medición, el autor del presente trabajo realiza la siguiente categorización de las métricas en (3) tres grupos, a fin de que sean tenidas en cuenta al momento de crear nuevos y diferentes indicadores a los ya propuestos y que probablemente puedan identificar durante su adopción:

- Con relación al Valor:
 - Contribución a los objetivos de negocio.
 - Reducción de riesgo por incumplimiento.
 - Mejora de la eficiencia en las operaciones.

- Con relación a la Eficacia:
 - Logro de metas y objetivos.
 - Utilizando las herramientas relevantes de gestión de datos.
 - Eficacia de la capacitación / formación.
 - Velocidad de adopción de cambio.

- Con relación a la Sustentabilidad:
 - Desempeño de las políticas y los procesos (es decir, ¿funcionan correctamente?)
 - Conformidad ¿Cumplen con los estándares y los procedimientos (es decir, ¿el personal sigue la orientación y cambia el comportamiento según sea necesario?).

A continuación, se plantean algunas métricas para el presente plan:

ID	Objetivo / Necesidad de Información	Métrica	Unidad	Fuente	Periodicidad del registro	Responsable	Cálculo
1	Reducir el porcentaje de registros duplicados	Valor	%	Se obtiene a partir del análisis de duplicados que se realiza con el proceso de calidad de datos.	Mensual	El equipo líder de gobierno de datos.	(Número de registros duplicados) / (Número de registros procesados)
2	Impacto de la capacitación	Eficacia	%	Promedio de la evaluación del Impacto de la capacitación en la eficacia de los Procesos	Mensual	Rol asignado	Evaluación del Impacto de la Capacitación en la Eficacia de los Procesos
3	Reducir el porcentaje de campos vacíos	Sustentabilidad	%	Se obtiene a partir del análisis de vacíos que se realiza con el proceso de calidad de datos.	Mensual	El equipo líder de gobierno de datos	(Número de campos vacíos) / (número de campos ingresados)

Tabla 20 Tabla de indicadores de medición propuestos
Fuente: Elaboración Propia

Las métricas para monitorear la calidad de los datos deben ser definidas por la Dirección de Apropiación bajo la orientación del equipo líder de gobierno de datos. En este caso, si la Dirección de Apropiación en el proceso de adopción de este plan, prioriza de todos sus conjuntos de datos 10 campos, el nivel de calidad deberá ser medido así:

- **Satisfactorio:** se considera satisfactorio cuando de los 10 los campos evaluados, entre el 90% y 100% cumple con las dimensiones.
- **Aceptable:** se considera aceptable cuando de los 10 campos evaluados, entre el 71% y 89% cumple con las dimensiones.

- **Inaceptable:** cuando en los 10 campos evaluados, entre el 0% y 70% cumple con las dimensiones.

90	hasta	100 %	■ Satisfactorio
71	hasta	89 %	■ Aceptable
0	hasta	70 %	■ Inaceptable

Figura 27 Métricas propuestas para medir la Calidad de Datos
Fuente: Elaboración Propia

Metas: dado que, al inicio de la implementación del plan, no se cuenta con una línea base de medición de calidad de datos a nivel de toda la Dirección de Apropiación, las metas deben ser establecidas una vez se realice el diagnóstico de las bases de datos y se conozca su respectivo nivel de calidad; así a partir de esta medición se definen las metas a alcanzar acorde con los indicadores establecidos en un periodo de tiempo.

5.5.6 Roles y responsabilidades

La matriz RACI, o matriz de responsabilidades, es una herramienta que permite identificar de forma esquemática la función de cada participante con relación a una determinada tarea o proceso (Luis Angulo, 2013). De esta forma se distinguen diferentes funciones:

- **R (Responsable):** es la persona que ejecuta la tarea.
- **A (Aprobador):** es la persona que debe aprobar el trabajo realizado y dar por concluida la tarea.
- **C (Consultado):** es la persona que presta ayuda al responsable.
- **I (Informado):** es la persona que debe estar informada de la ejecución de la tarea, pero sin participar de ella.

A continuación, se relaciona la matriz RACI propuesta para el presente plan; esta matriz se basa en el listado de procesos de gestión de datos identificados para la Dirección de Apropiación y de otras actividades que se identifican a continuación para la correcta adopción del plan propuesto.

ROLES Procesos / Actividades	Dirección de Apropiación Responsable Gestión de Datos	Oficina de TI	Coordinación de Contratación	Oficina de Planeación	Interesados en la Información	Operador de Servicio	Líder de Gobierno de Datos
Proceso Contractual	R	C	A	C	I	I	A
Proceso Captura de Datos	A	C	I	I		R	A
Proceso de Entrega al Ministerio	A	C	I	I		R	I
Proceso de Revisión	R	A		C		I	I
Proceso de Calidad	I	C		I			R
Proceso de Reporte de Información - Estandarización	R	C		I	A		A
Creación Procedimientos en el MIG	C	R		A	I		C
Plan de capacitación	A	R		C	I	I	A

Figura 28 Matriz de Roles y Responsabilidades RACI
Fuente: Elaboración Propia

5.5.7 Tecnología

Desde la Dirección de Gobierno Digital, se viene realizando la promoción y uso de herramientas tecnológicas que soporten las actividades de gestión de datos, conscientes de la necesidad de tomar decisiones basados en datos como uno de los propósitos del Marco de Referencia de Arquitectura Empresarial; es por ello, que acorde con esta promoción, se recomienda para la Dirección de Apropiación, el uso de la herramienta para calidad de datos llamada “Open Refine”.

Esta herramienta permite a los usuarios dejar de realizar los procesos de revisión y corrección de forma manual, para empezar a realizarlos de manera automática a través de esta herramienta; como es de código abierto, no implicará ningún costo adicional para su funcionamiento, adicionalmente es ideal para fuentes de datos provenientes de archivos de Excel, Csv, entre otros

Herramienta	Descripción	Licenciamiento
Open Refine.	<p>Open Refine, es una herramienta de código abierto, de fácil instalación, configuración y uso que permite a los funcionarios identificar y mejorar la calidad de los datos que gestionan en su quehacer diario o como OpenRefine (2018).</p> <p>Funcionalidad:</p> <ol style="list-style-type: none"> 1. Importar datos en varios formatos. 2. Explorar conjuntos de datos en cuestión de segundos. 3. Aplicar transformaciones celulares básicas y avanzadas. 4. Tratar con celdas que contienen múltiples valores. 5. Crear enlaces instantáneos entre conjuntos de datos. 6. Filtrar y particionar los datos fácilmente con expresiones regulares. 7. Utilizar la extracción de entidades con nombre en los campos de texto completo para identificar automáticamente los temas. 8. Realizar operaciones de datos avanzadas con el lenguaje general de expresión de refinamiento. 	Open Source.

*Tabla 21 Tabla de características de la Open Refine
Fuente: Elaboración Propia*

Si bien existen en el mercado una lista de herramientas para los mismos fines, se propone Open Refine, dado que es una herramienta libre y de uso sencillo es decir, las personas no requieren una capacitación exhaustiva y demasiado técnica para su utilización y obtención de los beneficios que requiere la Dirección, por lo tanto, se recomienda esta herramienta en su proceso de implementación del presente plan, no obstante, si la Dirección decide obtener cualquier otra herramienta, debe asesorarse con la Oficina de Tecnologías de Información.

5.5.8 Plan de implementación

Las actividades clave que la Dirección de Apropiación debe ejecutar para **generar cultura y sensibilización** sobre la importancia de la gestión de la calidad de datos y de los procedimientos sugeridos en el área que permita operacionalizar el modelo y el plan propuesto son:

- Realizar planes de capacitación y entrenamiento en conceptos básicos de gobierno de datos, a los equipos de trabajo que hacen parte de la estructura organizacional.
- Mejorar las habilidades y capacidades de gestión de datos del equipo de Negocio, Gobierno, Calidad y Protección de Datos, realizando capacitaciones técnicas en temáticas como:
 - Talleres para implementación del modelo de gobierno de datos.
 - Herramientas de calidad de datos, que permitan inicialmente automatizar el proceso de calidad de datos.
- Comunicar en la intranet del Ministerio mensajes significativos sobre temas de gestión de datos, como notas, artículos.
- Formalizar los roles y responsabilidades de los miembros del equipo líder de gobierno de datos.
- Involucrar a todos los interesados relevantes y fomentar su participación.
- Comunicar al interior de la entidad los miembros del equipo y datos de contacto.
- Comunicar el punto de contacto para requerimientos y/o necesidades de gestión de datos.

5.6 Recomendaciones

A continuación, se realiza un listado de recomendaciones que se pudieron encontrar en el ejercicio de la investigación, mediante las entrevistas de los expertos, aunado a esto sus recomendaciones fueron consideradas por el autor del presente trabajo acorde con su nivel de experiencia y conocimiento previo de la Entidad.

La gestión de cambio es también llamada capacitación para lograr cambios, y sirve para que las personas acepten los cambios que resultan de la implementación del modelo y para reducir los factores de rechazo. El proceso de cambio organizacional es complejo, es por esto que la Dirección de Apropiación, debe evitar a modo de recomendación incurrir en los siguientes errores al implementar la capacidad de gobierno de datos sobre la cual se acopla el plan de gestión de calidad:

La iniciativa no está liderada por TI, acorde con el experto entrevistado 4: la clave para el éxito de la gestión de datos consiste en lograr que las partes interesadas se responsabilicen de sus datos y apropien la iniciativa de gestión de datos como suya. A pesar de que en la oficina de TI no son propietarios de los datos, es significativo involucrarlos ya que entiende las implicaciones de no administrar los datos correctamente y, por lo tanto, son las personas que justifican que se necesita un gobierno de datos adecuado.

Entender el nivel de madurez actual del negocio, acorde con el experto entrevistado 4: hace referencia a que en muchas ocasiones no se entiende el nivel de madurez en términos de gestión de datos, es necesario por lo tanto que haya un entendimiento claro acerca de los procesos de gestión de datos que realiza la Dirección de Apropiación.

Gobierno de datos como un proyecto, acorde con el experto entrevistado 2: el secreto es implementar la iniciativa como un programa de cambio con diferentes flujos de trabajo, no como simplemente como un proyecto que tiene un inicio y fin, debe tener continuidad.

Alineación con la estrategia, acorde con el experto entrevistado 1: la Dirección de Apropiación debe tener claro por qué se está implementando esta iniciativa que hace parte del gobierno de datos y cómo esta puede ayudar a lograr los objetivos estratégicos. Es fundamental articular esto claramente para garantizar la financiación y el apoyo continuo.

Entender el estado de actual de los datos, acorde con el experto entrevistado 4: se debe tener un alto nivel de comprensión de cómo se administran los datos dentro del negocio, ya que actualmente se abordan los síntomas de falta de control de datos, en lugar de llegar a la causa raíz de estos.

Incorporar el marco propuesto, acorde con el experto entrevistado 2: los roles y las responsabilidades son una parte importante del modelo de gobierno de datos. El negocio debe asegurar que los roles y responsabilidades de gestión de datos estén bien definidas y que se seleccione las personas adecuadas para cada una de estas.

No pensar en que una herramienta es la respuesta, acorde con el experto entrevistado 1: el negocio antes de comprar una herramienta debe asegurarse de comprender completamente lo que está haciendo. Para aprovechar al máximo una herramienta, debe tener una comprensión clara de para qué la necesita, y si decide comprarla o implementar una de libre licenciamiento, debe recordar que estas herramientas no lo eximen de ninguna responsabilidad en términos de lograr un cambio organizacional.

6 Conclusiones

Desde el punto de vista metodológico de la investigación se concluye que, mediante su diseño, permitió de una manera más detallada la identificación de la pertinencia e importancia de los procesos de gestión de datos en las organizaciones y en específico en la Dirección de Apropiación. El enfoque descriptivo empleado y sus instrumentos de recolección permitieron evidenciar la problemática de manera más clara y objetiva a través de cada uno de los responsables de gestión de datos entrevistados, accediendo de manera más cercana a sus percepciones y reacciones frente al tema en cuestión; permitiendo enrutar al investigador hacia la identificación de otras causas que no habrían sido detectadas de otra manera.

La matriz de diagnóstico como soporte técnico a la justificación del problema, permitió evidenciar que las fuentes de información que gestiona la Dirección de Apropiación presentan niveles bajos en la calidad de sus datos, mediante los resultados del diagnóstico realizado, siendo esta la oportunidad de mejora más notable.

Como consecuencia del proceso de triangulación entre los marcos de referencia y las respuestas de los expertos, se logró de manera más sencilla obtener los criterios y buenas prácticas que debían ser parte como base para la construcción del plan como propuesta, dado a que éstos presentaban coincidencias que produjeron un nivel de confianza mayor.

En el análisis de los resultados de las entrevistas realizadas a los responsables de la gestión de datos, se pudieron identificar no solo las causas de la problemática, sino que, en el ejercicio de describir su situación, realizaban propuestas de mejora que desde el rol que ejercen han podido identificar que se pueden realizar, complementando de esta manera las recomendaciones que por su nivel de experiencia realizaban los expertos.

Todos y cada uno de los involucrados en el proceso de la investigación, lograron identificar los beneficios de la implementación de un plan de gestión de calidad de datos sin mayor esfuerzo, ya que acorde con el análisis de las entrevistas manifestaron la necesidad de una solución como la propuesta, que en el corto, mediano y largo plazo permita cerrar brechas en los procesos de gestión de datos en la Dirección de Apropiación mejorando los reportes de información y por ende la toma de decisiones.

Por último, en el plan de gestión de calidad sugerido se planteó no solo la importancia de la gestión del proceso de calidad, sino que tal y como lo destacaron los expertos, lo más importante en cualquier proyecto de datos, es el gobierno de datos, dado que si éste no existe, es aún más probable que los proyectos estén

destinados al fracaso; por esa razón, el plan propone un modelo de gobierno de datos que lleva implícito la práctica de calidad de datos, convirtiéndose en una herramienta que unifica los conceptos a considerar por parte de la Entidad para el mejoramiento de sus datos a través de los procesos.

Referencias Bibliográficas

- AIIM International. (2009). *Hello Mike 2.0*. 1–3.
- Alvarado, J. V. (2011). *Calidad de datos en las organizaciones*. 68.
- Birnbaum, M. L. (2007). Data, information, and indicators. *Prehospital and Disaster Medicine*, Vol. 22, pp. 471–472. <https://doi.org/10.1017/S1049023X00005276>
- Bonilla-Castro, E. & Rodriguez Sehk, P. (2005). *Más allá del dilema de los métodos: la investigación en ciencias sociales* (Ediciones Uniandes, Ed.). Retrieved from https://books.google.com.co/books?id=REOIWoQuAL4C&pg=PA138&dq=muestreo+teórico&hl=es-419&sa=X&ved=0ahUKEwi50qX_qIPmAhWp1FkKHbLUD7UQ6AEIMTAB#v=onepage&q=muestreo+teórico&f=false
- Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14. <https://doi.org/10.5334/dsj-2015-002>
- Calero, C., Moraga, M. Á., & Piattini, M. (2010). Calidad Del Producto Y Proceso Software. In E. Rama (Ed.), *DFVCL*. <https://doi.org/ISBN:8478979611> ISBN-13: 9788478979615
- Congreso de la República. *Ley 1712*. , (2014).
- Consejo Nacional de Política Económica y Social [CONPES]. (2018). *Política Nacional de Explotación de Datos, Documento CONPES 3920*. <https://doi.org/https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3920.pdf>
- Corporación Colombia Digital [CCD]. (2017). Beneficios de una buena calidad de datos y cómo perjudica si es mala. Retrieved October 9, 2019, from Octubre 13 website: <https://colombiadigital.net/actualidad/articulos-informativos/item/9872-beneficios-de-una-buena-calidad-de-datos-y-como-perjudica-si-es-mala.html>
- DAMA. (2008). *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK) Spanish Edition - Google Play* (2010th ed.; Technics Publications, Ed.). Retrieved from <https://books.google.com.co/books?id=7FdNCwAAQBAJ&pg=PT18&lpg=PT18&dq=las+organizaciones+que+no+entienden+la+gran+importancia+de+la+gestión+de+datos+y+la+información+como+activos+tangibles+en+la+nueva+economía+no+van+a+sobrevivir&source=bl&ots=9uflO1m79P&>
- DAMA. (2017). Body of Knowledge DAMA. Retrieved October 9, 2019, from <https://dama.org/content/body-knowledge>
- Departamento Nacional de Estadística [DANE]. (n.d.). *Plan de Fortalecimiento de Registros Administrativos: Estrategia para el Fortalecimiento Estadístico Territorial*.
- Departamento Nacional de Planeación [DNP]. (2019a). El Consejo Nacional de Política Económica y Social, CONPES. Retrieved October 9, 2019, from <https://www.dnp.gov.co/CONPES/Paginas/conpes.aspx>

- Departamento Nacional de Planeación [DNP]. (2019b). La transformación digital como generador transversal de valor económico y social: la visión del PND. Retrieved October 27, 2019, from 6 de Septiembre de 2019 website: <https://www.dnp.gov.co/Paginas/La-transformacion-digital-como-generador-transversal-de-valor-economico-y-social-la-vision-del-PND.aspx>
- Deusto Formación. (2018). 4 causas de mala calidad de los datos en Big Data. Retrieved October 9, 2019, from 12/10/2018 website: <https://www.deustoformacion.com/blog/blog-empresa-nuevas-tecnologias/4-causas-mala-calidad-datos-big-data>
- Discover The New. (2017). El alto coste de la mala calidad de los datos. *19 DIC 2017*. Retrieved from <https://discoverthenew.ituser.es/predictive-analytics/2017/12/el-alto-coste-de-la-mala-calidad-de-los-datos>
- DOMO. (2019). Data Never Sleeps 7.0: Infographic. Retrieved October 9, 2019, from https://www.domo.com/learn/data-never-sleeps-7?utm_source=wire&utm_medium=pr&utm_campaign=ABM_Other_FY20_Global_PR&campid=701f2000001C3kNAAS
- Durán, D. (2019a). *Instrumentos de investigación cualitativos y cuantitativos frente a la investigación mixta o complementaria*. 3, 41–56. Retrieved from <http://pragmatika.cl/review/index.php/consensus/article/download/38/50/>
- Durán, D. (2019b). *Instrumentos de investigación cualitativos y cuantitativos frente a la investigación mixta o complementaria*. 3, 41–56. Retrieved from <http://pragmatika.cl/review/index.php/consensus/article/download/38/50/>
- Echegoyen, G. (2003). *Registros administrativos, calidad de los datos y credibilidad pública: presentación y debate de los temas sustantivos de la segunda reunión de la Conferencia Estadística de las Américas de la CEPAL*. Retrieved from <http://www.cepal.org/publicaciones/xml/8/14038/lc2024e.pdf>
- Fisher, T. (2009). *El activo de datos: cómo las empresas inteligentes gobiernan sus datos para el éxito empresarial*. Retrieved from [https://books.google.es/books?hl=es&lr=&id=ZplvI7Q2C2UC&oi=fnd&pg=PR5&dq=Fisher,+T.++\(2009\).+The+Data+Asset.+How+Smart+Companies+Govern+Their+Data+for+Business+Success.+Estados+Unidos:+John+Wiley+%26+Sons,+Inc.&ots=dX-gYCBu93&sig=JwXZxxMQk2_kOZRIHwmPqmFqff](https://books.google.es/books?hl=es&lr=&id=ZplvI7Q2C2UC&oi=fnd&pg=PR5&dq=Fisher,+T.++(2009).+The+Data+Asset.+How+Smart+Companies+Govern+Their+Data+for+Business+Success.+Estados+Unidos:+John+Wiley+%26+Sons,+Inc.&ots=dX-gYCBu93&sig=JwXZxxMQk2_kOZRIHwmPqmFqff)
- fluxicon. (n.d.). *Lista de verificación de calidad de datos para Procesos de Minería*. (0).
- Gati Consultores. (2019). Gobierno de Datos Inteligente. *Conferencia Sobre Gestión de Datos*. Bogotá D.C.
- Gobierno Digital. (n.d.). Gobierno Digital. Retrieved October 9, 2019, from <https://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7650.html>
- Hernández Sampieri, R., Collado, C. F., & Baptista Lucio, M. del P. (2010). *Metodología de la investigación, 5ta Ed*. Retrieved from www.FreeLibros.com
- IBM. (2012). Seis pasos para el Gobierno de Datos: ¿Qué es y cómo se implementa un programa de Gobierno de Datos? Retrieved November 26, 2019, from 23-07-2012 website:

- <https://www.ibm.com/developerworks/ssa/data/library/techarticle/gobierno-datos/>
InfoChannel. (2017). DAMA llega a México para dar valor a los datos. *Lun, 20/02/2017 - 12:34*. Retrieved from <https://www.infochannel.info/dama-llega-mexico-para-dar-valor-los-datos>
- International Dynamic Advisors [Intedy]. (n.d.). *¿Qué son las ISO? ¿Qué es ISO? ¿Cuál fue la primera norma ISO publicada?* Retrieved from <https://www.intedy.com/internacional/intedy-noticias.php?id=1589#submenuhome>
- International Organization for Standardization [ISO]. (n.d.). ISO 9000 family – Quality management. Retrieved October 9, 2019, from <https://www.iso.org/iso-9001-quality-management.html>
- ISO 8000. (2018). Normas ISO 8000. Retrieved October 20, 2019, from <http://iso8000.es/normas-iso-8000>
- ISOTools. (2015). ¿Qué son las normas ISO y cuál es su finalidad? Retrieved October 9, 2019, from 19 MARZO, 2015 website: <https://www.isotools.org/2015/03/19/que-son-las-normas-iso-y-cual-es-su-finalidad/>
- Jager, S. De. (2016). *ISO 8000 MD Introduction*.
- Keller, S., Korkmaz, G., Orr, M., Schroeder, A., & Shipp, S. (2017). The Evolution of Data Quality: Understanding the Transdisciplinary Origins of Data Quality Concepts and Approaches. *Annual Review of Statistics and Its Application, 4(1)*, 85–108. <https://doi.org/10.1146/annurev-statistics-060116-054114>
- Kimball, R. (2004). *Consejo de diseño de Kimball # 59: valor sorprendente del perfil de datos*. (59), 1–2. Retrieved from <http://www.kimballgroup.com/wp-content/uploads/2012/05/DT59SurprisingValue.pdf>
- Loshin, D. (2001). Gestión del conocimiento empresarial: el enfoque de calidad de datos. Retrieved November 26, 2019, from [https://books.google.com.co/books?id=3BXTfCtR8zsC&pg=PA73&dq=Enterprise+Knowledge+Management+en+español&hl=es-419&sa=X&ved=0ahUKEwjSqK67vYjmAhXyw1kKHXMCCUQ6AEIKDAA#v=onepage&q=Enterprise Knowledge Management en español&f=false](https://books.google.com.co/books?id=3BXTfCtR8zsC&pg=PA73&dq=Enterprise+Knowledge+Management+en+español&hl=es-419&sa=X&ved=0ahUKEwjSqK67vYjmAhXyw1kKHXMCCUQ6AEIKDAA#v=onepage&q=Enterprise+Knowledge+Management+en+español&f=false)
- Luis Angulo. (2013). *Gestión de Proyectos con Project, Excel y Visio*. Retrieved from https://books.google.com.co/books?id=cbkuDgAAQBAJ&pg=PA96&dq=matriz+raci&hl=es-419&sa=X&ved=0ahUKEwikvK_xwLrIAhUpwFkKHWJzB1UQ6AEINjAC#v=onepage&q=matriz+raci&f=false
- Mar, B. (2018). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Retrieved October 9, 2019, from Forbes website: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#658593b560ba>
- Méndez del Río, L. (2006). *Más allá del Business Intelligence: 16 experiencias de éxito* (Grupo Planeta (GBS), Ed.). Retrieved from <https://books.google.com.co/books?id=6ciGp2-NYewC&pg=PA30&dq=calidad+de+datos&hl=es->

419&sa=X&ved=0ahUKEwiwx6qi0ebiAhVBmVvKkHYCrA5QQ6AEIPTAE#v=onepage&q&f=false

- MIKE 2.0. (n.d.-a). MIKE2.0 - Data and Social Methodology, Open Framework | Information Management - MIKE2.0, the open source methodology for Information Development. Retrieved October 9, 2019, from http://mike2.openmethodology.org/wiki/What_is_MIKE2.0
- MIKE 2.0. (n.d.-b). Mike2 core solution offerings.jpg. Retrieved October 20, 2019, from http://mike2.openmethodology.org/wiki/File:Mike2_core_solution_offerings.jpg
- [MinTIC], M. de T. de la I. y las C. (n.d.). Marco de Referencia Anterior - Arquitectura TI. Retrieved November 28, 2019, from <https://www.mintic.gov.co/arquitecturati/630/w3-propertyvalue-8114.html>
- [MinTIC], M. de T. de la I. y las C. *Resolución 0001998.pdf.*, (2017).
- [MinTIC], M. de T. de la I. y las C. (2019a). Organigrama - Ministerio de Tecnologías de la Información y las Comunicaciones. Retrieved November 28, 2019, from <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/Organigrama/>
- [MinTIC], M. de T. de la I. y las C. (2019b). Organigrama - Ministerio de Tecnologías de la Información y las Comunicaciones. Retrieved November 28, 2019, from <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/Organigrama/>
- [MinTIC], M. de T. de la I. y las C. (2019c). Organigrama - Ministerio de Tecnologías de la Información y las Comunicaciones. Retrieved November 28, 2019, from <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/Organigrama/>
- [MinTIC], M. de T. de la I. y las C. (2019d). Organigrama - Ministerio de Tecnologías de la Información y las Comunicaciones. Retrieved November 28, 2019, from <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/Organigrama/>
- [MinTIC], M. de T. de la I. y las C. (2019e). Organigrama - Ministerio de Tecnologías de la Información y las Comunicaciones. Retrieved November 28, 2019, from <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/Organigrama/>
- Ministerio de Industria Comercio y Turismo. (2013). *Decreto 1375 de 27 Junio 2013* (pp. 1–8). pp. 1–8. Retrieved from https://www.mintic.gov.co/portal/604/articles-4274_documento.pdf
- Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC]. (2014). Instituciones relacionadas. Retrieved October 11, 2019, from 12 de febrero de 2014 website: <http://www.santpau.cat/web/public/institucions-relacionades>
- Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC]. (2019). *Manual de Gobierno Digital: Implementación de la Política de Gobierno Digital*. Retrieved from https://estrategia.gobiernoenlinea.gov.co/623/articles-81473_recurso_1.pdf
- MinTIC. (2015). *Decreto Único Sector TIC*. Retrieved from <https://www.mintic.gov.co/portal/604/w3-propertyvalue-14797.html>
- MinTIC. (2018). Contexto - Arquitectura TI. Retrieved October 9, 2019, from <https://www.mintic.gov.co/arquitecturati/630/w3-propertyvalue-8109.html>

- MinTIC. (2019a). Acerca del MinTIC. Retrieved October 9, 2019, from 22 de agosto de 2019 website: <https://www.mintic.gov.co/portal/inicio/Ministerio/Acerca-del-MinTIC/>
- MinTIC. (2019b). Dirección de Apropiación de TIC. Retrieved October 9, 2019, from 01 de marzo de 2019 website: <https://www.mintic.gov.co/portal/inicio/Ministerio/Viceministerio-de-Economia-Digital/Direccion-de-Apropiacion-de-TIC/>
- MinTIC. (2019c). *Estrategia de Integración Digital del Estado: Lineamientos para la integración de entidades públicas al Portal Único del Estado Colombiano - GOV.CO.*
- MinTIC. (2019d). Oficina de Tecnologías de la Información. Retrieved October 9, 2019, from 25 de enero de 2019 website: <https://www.mintic.gov.co/portal/inicio/Ministerio/Ministra/Oficina-de-Tecnologias-de-la-Informacion/>
- Mundo BI. (2011). MIKE 2.0 – Gestión de información « Mundo BI. Retrieved October 20, 2019, from mayo 17, 2011 website: <http://mundobi.com.ar/?p=322>
- Nogales, Á. F. (2004). *Investigación y Técnicas de Mercado* (2a Edición; ESIC Editorial, Ed.). Retrieved from <https://books.google.com.co/books?id=LnVxgMkEhkgC&pg=PA154&dq=tamaño+de+la+muestra+por+conveniencia&hl=es-419&sa=X&ved=0ahUKEwjh1ci95-zjAhWktlkKHTxJCtkQ6AEIKDAA#v=onepage&q=tamaño+de+la+muestra+por+conveniencia&f=false>
- OpenRefine. (2018). [openrefine.github.com](http://openrefine.org/). Retrieved October 20, 2019, from <http://openrefine.org/>
- Pipino, L. L., Lee, Y. W., Wang, R. Y., & Yang, R. Y. (2002). *Data Quality Assessment* (Vol. 45).
- PowerData. (n.d.). *La calidad de los datos : una radiografía completa*. 1–41.
- PowerData. (2018). Transformación digital. Qué es y su importancia y relación con los datos. Retrieved October 27, 2019, from PowerData website: <https://www.powerdata.es/transformacion-digital>
- Quintanilla Cobián, Laura; García Gallego, Carmen; Rodríguez Fernández, Raquel; Fontes de Gracia, Sofía, Sarriá Sánchez, E. (2019). *Fundamentos de investigación en Psicología* (Editorial UNED, Ed.). Retrieved from https://books.google.com.co/books?id=_z6wDwAAQBAJ&pg=PT125&dq=muestreo+por+conveniencia+validéz&hl=es-419&sa=X&ved=0ahUKEwiD-MCWtIPmAhXFslkKHfDbAv4Q6AEINjAC#v=onepage&q=muestreo+por+conveniencia+validéz&f=false
- República de Colombia. (2019). *DIRECTIVA PRESIDENCIAL No. 02.*
- Santos, M. P., & Lucas, A. (2019). *Identificación de factores críticos de éxito para la gestión de la calidad de los datos a través de un estudio de Delphi*. 13(8), 437–442.
- SAS Latin America. (2016). La importancia de una gestión de datos adecuada en las organizaciones. Retrieved October 11, 2019, from <https://blogs.sas.com/content/sasla/2016/05/12/la->

importancia-de-una-gestion-de-datos-adeuada-en-las-organizaciones/

- Seid, G. (2016). *Procedimientos para el análisis cualitativo de entrevistas. Una propuesta didáctica. V Encuentro Latinoamericano de Metodología de las Ciencias Sociales*. Retrieved from <http://elmeccs.fahce.unlp.edu.ar>
- Silva, M. R. F. da. (2006). *Introducción a las técnicas cualitativas de investigación aplicadas en salud*. Retrieved from <https://books.google.com.co/books?id=o2n57QYwMDIC&pg=PA44&dq=muestreo+teórico&hl=es-419&sa=X&ved=0ahUKEwiDjInX4ezjAhWy1IkKHf0PCyQQ6AEIPDAD#v=onepage&q=muestreo+teórico&f=false>
- Thomas, G. (n.d.). Definición de gobernanza de datos - The Data Governance Institute. Retrieved November 26, 2019, from <http://www.datagovernance.com/defining-data-governance/>
- Velthuis, M. G. P. (2018). *Calidad de Datos* (RA-MA Editorial, Ed.). Retrieved from https://books.google.com.co/books?id=y16fDwAAQBAJ&pg=PA26&lpg=PA26&dq=datos+malos+hacen+imposible+una+buena+toma+de+decisiones&source=bl&ots=jfaM4WJy22&sig=ACfU3U114pJZXMM_dQUVUhmWg0D7741-JA&hl=es-419&sa=X&ved=2ahUKEwjF8qChqpTIAhUliOAKHRixBf8Q6AEwAHoECAkQ
- Wade, M., Loucks, J., Macaulay, J., & Noronha, A. (2018). *Digital Vortex* (L. Editorial, Ed.). Retrieved from https://books.google.com.co/books?id=EcpsDwAAQBAJ&printsec=frontcover&dq=How+Today%27s+Market+Leaders+Can+Beat+Disruptive+Competitors+at+Their+Own+Game.&hl=es-419&sa=X&ved=0ahUKEwi0mImzturiAhUo2FkKHd_DnoQ6wEIKTAA#v=onepage&q=How+Today's+Market+Leaders+Ca
- Wallgren, A., & Wallgren, B. (2016). *Hacia un sistema estadístico integrado y basado en registros*.
- Westerman, G., Bonnet, D., & McAfee, A. (2014). Leading Digital. *Harvard Business Review Press*, 1–292. Retrieved from <https://www.lehmanns.de/shop/wirtschaft/30813547-9781625272485-leading-digital>
- White, D. (2019, March 21). Por qué los datos serán la clave del éxito en el futuro de las empresas. *BBVA: Opinión*. Retrieved from <https://www.bbva.com/es/opinion/por-que-los-datos-seran-la-clave-del-exito-en-el-futuro-de-las-empresas/>

Anexos

Anexo 1 Guía I de entrevista a responsables de gestión de datos.

UNIVERSIDAD EXTERNADO DE COLOMBIA GUÍA I DE ENTREVISTA SEMIESTRUCTURADA A RESPONSABLES DEL MANEJO DE DATOS EN DIRECCIÓN DE APROPIACIÓN DEL MINTIC	
MÁGISTER EN GERENCIA ESTRATÉGICA DE TI	
PROYECTO: “PLAN DE GESTIÓN DE CALIDAD DE DATOS PARA MEJORAR LA OPORTUNIDAD Y PERTINENCIA DE LA INFORMACIÓN DE LA OFERTA INSTITUCIONAL EN EL MINISTERIO DE TECNOLOGÍAS DE LA INFORMACIÓN.	
Fecha de entrevista:	
Profesión:	Cargo:
<u>Introducción</u> La presente entrevista tiene como fin, describir la problemática expuesta en el trabajo de grado.	
Preguntas CONOCIMIENTO DEL PROCESO DE LA GESTIÓN DE DATOS Y PROBLEMAS DETECTADOS. 1. Por favor, descríbame el proceso que se realiza en cuanto a la gestión de datos, desde su recepción hasta el reporte a las áreas solicitantes. 2. ¿Desde el rol que desempeña en temas de Gestión de Datos, ha tenido dificultades con los datos que reportan los operadores de servicio? ¿cuales? 3. ¿Por qué los datos que reportan a las diferentes áreas de la Entidad están presentando inconsistencias tales como: -Falta de integridad en los datos. -Falta de oportunidad. -Campos vacíos. -Duplicidad en los registros. -Tipos de datos no conformes.	

4. ¿Sabe usted cómo resolver alguno de los problemas mencionados en la pregunta anterior?
5. ¿Desde su función, qué tipo de inconvenientes ha tenido con otras áreas al momento de reportar los datos de la Dirección de Apropiación? ¿Con qué frecuencia ocurren estos inconvenientes?
6. ¿Cómo realiza la revisión de los datos que le entregan los operadores de servicio a la Dirección de Apropiación?

PROCEDIMIENTOS ACTUALES.

7. ¿Dentro del proceso diario que realiza, ha escuchado alguna vez por parte de los operadores de servicio de los proyectos o dentro de su área de trabajo el concepto de Calidad de Datos?
8. ¿Sabe usted a qué se refiere ese concepto (Calidad de Datos)?
9. ¿Durante el proceso contractual, cuáles son las condiciones mínimas que se exigen a los operadores de servicio para el reporte de información al Ministerio?
10. ¿Qué métodos o herramientas utiliza para medir las diferentes bases de datos que gestiona desde su cargo?
11. ¿Conoce usted algún tipo de proceso ya sea formal o no, para la gestión de datos?
12. ¿De acuerdo con su experiencia en el cargo y desde sus funciones tiene en cuenta manuales, guías, hojas de ruta o procedimientos del MinTIC para la gestión de datos?
13. ¿Cuáles procedimientos o guías toma como referencia la Dirección de Apropiación para el manejo y control de los datos?
14. ¿Sabe usted a quién dirigirse en caso de tener problemas en el proceso de gestión de datos?

CONSIDERACIONES DE PROPONER UN PLAN DE GESTIÓN

15. ¿Considera usted, de acuerdo con su experiencia en el área, que hay necesidades por atender en cuanto al control de la captura, recepción y gestión de los datos que proporcionan los operadores de servicio a la Dirección de Apropiación?

16. Dé su opinión sobre la siguiente afirmación <La Dirección de Apropiación presenta la necesidad de implementar o adoptar alguna estrategia que le permita mejorar sus procesos de reporte de datos>
17. ¿Considera usted que es importante implementar un plan de gestión de datos adoptado a la Dirección de Apropiación? ¿Por qué?
18. ¿Si se adopta un plan de gestión de datos y/o algún procedimiento, considera que podrían mejorar los procesos de reporte de información en el área? ¿Por qué?
19. ¿Considera que la adopción de un plan de gestión de calidad de datos impactaría de manera positiva o negativa en los recursos financieros y de tiempo hacia los responsables de gestionar los datos en el área?

Anexo 2 Guía II de entrevista a expertos en gestión de datos.

UNIVERSIDAD EXTERNADO DE COLOMBIA GUÍA II ENTREVISTA SEMIESTRUCTURADA A EXPERTOS	
MÁGISTER EN GERENCIA ESTRATÉGICA DE TI	
PROYECTO: “PLAN DE GESTIÓN DE CALIDAD DE DATOS PARA MEJORAR LA OPORTUNIDAD Y PERTINENCIA DE LA INFORMACIÓN DE LA OFERTA INSTITUCIONAL EN EL MINISTERIO DE TECNOLOGÍAS DE LA INFORMACIÓN.	
Fecha de entrevista:	
Profesión:	Empresa en la que labora:
Cargo o Rol:	Tiempo de Experiencia Profesional:
<u>Introducción</u> La presente entrevista tiene como fin, identificar desde una perspectiva profesional los criterios y buenas prácticas que existen para la implementación de un plan de gestión de calidad de datos.	
Preguntas INTRODUCCIÓN AL PROBLEMA DESDE LA PERSPECTIVA DE LOS EXPERTOS. <ol style="list-style-type: none">1. En su trayectoria en la gestión de datos ¿qué problemas son comunes encontrarse en el camino?2. De acuerdo con su experiencia, ¿qué tan común en las organizaciones, es el problema de la falta de calidad en los datos y a que se debe? ENTORNO IDEAL PARA LA GESTIÓN DE LOS DATOS <ol style="list-style-type: none">3. ¿Qué criterios y o mecanismos de control de calidad de los datos considera, que deben tener en cuenta las áreas de una organización, cuyas fuentes no provienen directamente de sistemas de información?4. ¿Considera usted, que, para los procesos de gestión de datos, las personas encargadas deben tener conocimientos acerca de la labor que realizan?5. ¿Qué tipo de conocimientos básicos deberían tener las personas encargadas de la captura y gestión de los datos, sin ser profesionales en TI?	

EL PLAN DE GESTIÓN DE CALIDAD DE DATOS COMO HERRAMIENTA DE SOLUCIÓN.

6. ¿Considera usted importante contar con una metodología o un plan o algún otro tipo de instrumento para la gestión de los datos en las organizaciones? ¿por qué?
7. ¿Considera que la implementación de un plan de gestión de calidad de datos mejoraría notablemente los procesos de reporte de información en las áreas? ¿Por qué?
8. ¿Cuál es la razón principal por la que las organizaciones deben implementar planes de calidad de datos?
9. ¿Qué beneficios se podrían obtener en la adopción de un plan de calidad, de cara a los involucrados en los procesos de gestión de datos?
10. Recomiende 4 puntos clave que debería tener un buen plan de gestión de calidad de datos.

MARCOS DE REFERENCIA

11. ¿Tiene conocimiento acerca de lo que es el DAMA?
12. ¿Cuáles marcos de referencia que usted considera se deben tener en cuenta para la implementación de un plan de gestión de calidad de datos?
13. ¿Conoce usted la metodología MIKE 2.0?
14. Dado que las ISO, son estándares internacionales y el DAMA es reconocido internacionalmente también, ¿cuál considera que es un mejor referente para ser adaptado a las necesidades de la entidad?