

MARCO DE REFERENCIA PARA UN MODELO DE GOBIERNO DE DATOS

-Caso de uso en una Caja de Compensación Familiar-

PABLO ANDRES TOBAR RUIZ

**Escuela Colombiana de Ingeniería Julio Garavito
Decanatura de Ingeniería de Sistemas
Maestría Gestión de Información
Bogotá D.C., 25 de enero de 2021**



MARCO DE REFERENCIA PARA UN MODELO DE GOBIERNO DE DATOS

-Caso de uso en una Caja de Compensación Familiar-

PABLO ANDRES TOBAR RUIZ

**Trabajo de investigación para optar al título de
Magíster en Gestión de Información**

Directores

**VICTORIA EUGENIA OSPINA BECERRA
PhD.**

**DANIELA CAROLINA DE LA ROSA PÉREZ
Ing.**

**Escuela Colombiana de Ingeniería Julio Garavito
Decanatura de Ingeniería de Sistemas
Maestría en Gestión de Información
Bogotá D.C., 25 de enero de 2021**



PÁGINA DE ACEPTACIÓN

El trabajo de grado de maestría titulado “MARCO DE REFERENCIA PARA UN MODELO DE GOBIERNO DE DATOS, CASO DE USO EN UNA CAJA DE COMPENSACIÓN FAMILIAR”, presentado por PABLO ANDRÉS TOBAR RUIZ, cumple con los requisitos establecidos y recibe nota aprobatoria para optar al título de Magíster en Gestión de información.



NOMBRE DIRECTOR 1
Director del Trabajo de Grado



NOMBRE DIRECTOR 2
Director del Trabajo de Grado



NOMBRE JURADO 1
Jurado



NOMBRE JURADO 2
Jurado

Bogotá, D.C., 26 de ENERO de 2021

© Únicamente se puede usar el contenido de las publicaciones para propósitos de información. No se debe copiar, enviar, recortar, transmitir o redistribuir este material para propósitos comerciales sin la autorización de la Escuela Colombiana de Ingeniería. Cuando se use el material de la Escuela se debe incluir la siguiente nota "Derechos reservados a Escuela Colombiana de Ingeniería" en cualquier copia en un lugar visible. Y el material no se debe notificar sin el permiso de la Escuela.

Publicado en 2020 por la Escuela Colombiana de Ingeniería Julio Garavito. Avenida 13 No 205-59 Bogotá. Colombia TEL: +57 – 1 668 36 00

Resumen

En un entorno empresarial y social donde los datos se convirtieron en el factor diferenciador entre el éxito o el fracaso, ha surgido la necesidad de implementar nuevas capacidades que permitan de manera precisa, ejercer control y gobierno sobre los datos en todo su ciclo de vida.

El gobierno que se ejerza sobre los datos desde el punto de vista estratégico, técnico y operativo, deberá ser con el tiempo un proceso inherente al funcionamiento de todo tipo de organización, integrándose dentro del quehacer de los negocios y de la sociedad en general.

Estas metodologías emergentes en torno a los datos y a la información, requieren de un proceso de entendimiento y análisis previo para lograr una ejecución exitosa; por tal motivo, en este documento se propone un marco de referencia para entender y ejecutar un programa de gobierno de datos, basados en la experiencia de una empresa del sector de la seguridad social en Colombia.

Se abordan elementos que permitirán desarrollar el proceso de implementación de un modelo de gobierno de datos, empezando por el análisis de situación actual para visionar la situación futura; y de esta manera lograr tomar decisiones basados en una gestión adecuada de datos.

Palabras clave: gestión de datos, explotación de datos, gobierno de datos, cajas de compensación.

Abstract

In a social environment where data became the differentiating factor between business success or failure, there has been a need to implement new capabilities that allow precise, exercise control and governance over the data throughout its life cycle.

The government that exerts on the data from the strategic, technical and operational point of view, will have to be over time a process inherent to the operation of all types of organizations, integrating into the business and society in general.

These emerging methodologies around data and information require a process of understanding and prior analysis to achieve a successful implementation; for this reason, this document proposes a framework to understand and implement a data governance program, based on the experience of a company not profit in Colombia.

Different elements are addressed that will allow to develop the process of implementing a data governance model, starting with the analysis of the current situation to view the future situation; and in this way achieve decisions based on proper data management.

Keywords: data management, data use, data governance, not-profit company

Índice General

1	INTRODUCCIÓN	9
1.1	JUSTIFICACIÓN	9
1.2	OBJETIVOS	10
1.2.1	<i>General</i>	10
1.2.2	<i>Específicos</i>	10
1.3	ALCANCE Y LIMITACIONES	10
1.4	METODOLOGÍA	11
2	MARCO CONTEXTUAL Y CONCEPTUAL	12
2.1	CONCEPTOS BÁSICOS	12
2.2	SOBRE LOS DATOS	12
2.3	SOBRE EL GOBIERNO DE DATOS	14
2.4	SOBRE LA GESTIÓN DE DATOS	19
2.5	SOBRE LAS CAJAS DE COMPENSACIÓN FAMILIAR	21
2.6	DEL ESTADO DEL ARTE	24
3	DIAGNÓSTICO	28
3.1	MODELO DE MADUREZ	28
3.2	DESARROLLO DEL CASO PRÁCTICO	29
3.3	DEFINICIÓN DEL NIVEL DE MADUREZ	30
3.4	OPORTUNIDADES E INDUCTORES	32
4	MARCO DE REFERENCIA	34
4.1	FACTORES DE ÉXITO	35
	<i>Misión</i>	35
	<i>Visión</i>	35
	<i>Objetivos</i>	36
	<i>Alcance</i>	36
	<i>Impulsores</i>	36
	<i>Capacidades</i>	37
	<i>Políticas</i>	41
4.2	MODELO DE OPERACIÓN	42
4.2.1	<i>De la gestión operativa de datos</i>	42
4.2.2	<i>Del comité de gobierno de datos</i>	45
4.2.3	<i>De la estructura organizacional</i>	46
4.3	CADENA DE VALOR	51
4.3.1	<i>De la organización basada en datos</i>	51
4.3.2	<i>Del comité de gobierno de datos</i>	53
4.3.3	<i>De la explotación y exploración de datos</i>	55
4.4	CICLO DE VIDA DE LOS DATOS	56

4.5	LINEAMIENTOS, DISCIPLINAS, Y/O ÁREAS DE CONOCIMIENTO.....	58
4.5.1	<i>Lineamientos generales</i>	59
4.5.2	<i>Disciplinas principales</i>	65
4.5.3	<i>Disciplinas de soporte</i>	79
4.6	CULTURA Y CAMBIO ORGANIZACIONAL.....	88
4.6.1	<i>Objetivos</i>	88
4.6.2	<i>Acciones</i>	88
4.6.3	<i>Indicadores de seguimiento</i>	89
5	RECOMENDACIONES	91
5.1	MAPA DE RUTA.....	91
5.2	INDICADORES DE SEGUIMIENTO Y CONTROL.....	93
6	CONCLUSIONES	95
6.1	TRABAJOS FUTUROS.....	96
	REFERENCIAS BIBLIOGRÁFICAS	98
	ANEXOS	100
	ANEXO 1. ENCUESTA DE ANÁLISIS EXPLORATORIO Y DE MEDICIÓN DE PERCEPCIÓN.....	100
	ANEXO 2. ENCUESTA PERCEPCIÓN PARA IDENTIFICAR ADHERENCIA DEL PROGRAMA DE GOBIERNO DE DATOS.....	106
	ANEXO 3. DECÁLOGO DEL GOBIERNO DE DATOS.....	106

Índice de ilustraciones

Ilustración 1 Modelo alineación estratégica.....	15
Ilustración 2 Modelo de la información de Amsterdam.....	16
Ilustración 3 Pirámide de Peter Aiken	17
Ilustración 4 Propuesta gobierno de datos IBM.....	18
Ilustración 5 Estructura de gobierno de datos DGI.....	19
Ilustración 6 Áreas de conocimiento.....	20
Ilustración 7 Diagrama de contexto	21
Ilustración 8 Factores ambientales.....	21
Ilustración 9 Diagrama Canvas para una CCF	23
Ilustración 10 Ámbitos del gobierno del dato.....	25
Ilustración 11 Valoración del nivel de madurez	31
Ilustración 12 Pilares del modelo de gobierno de datos	37
Ilustración 13 Operación sistemas transaccionales CCF	43
Ilustración 14 Gestión operativa de datos	43
Ilustración 15 Modelo operativo del comité de gobierno de datos.....	45
Ilustración 16 Estructura organizacional nivel corporativo.....	47
Ilustración 17 Estructura organizacional a nivel operativo de negocio	50
Ilustración 18 Modelo de una organización manejada por datos.....	52
Ilustración 19 Macro-procesos del comité de gobierno de datos.....	53
Ilustración 20 Cadena de valor - explotación y exploración de datos.....	55
Ilustración 21 Modelo del ciclo de vida de los datos.....	56
Ilustración 22 Modelo de gobierno de datos desde las disciplinas o áreas de conocimiento.....	59
Ilustración 23 Modelo dimensional para la entidad afiliado	87
Ilustración 24 Modelo conceptual para la entidad afiliado	87

Índice de tablas

Tabla 1 Identificación inicial de procesos	38
Tabla 2 Formato levantamiento de procesos	38
Tabla 3 Identificación inicial de artefactos tecnológicos	39
Tabla 4 Identificación inicial de áreas y personas	40
Tabla 5 Fases de operación de políticas	41
Tabla 6 Matriz de priorización de iniciativas de datos	54
Tabla 7 Ciclo de vida para la entidad de afiliado	58
Tabla 8 Ciclo de vida para asignación de crédito	58
Tabla 9 Ejemplo glosario de negocio	63
Tabla 10 Identificación de entidades y sub-entidades para una CCF	64
Tabla 11 Calidad de datos: atributos, dimensiones, reglas, roles y responsabilidades	69
Tabla 12 Calidad de datos: dimensiones y criterios de negocio	69
Tabla 13 Calidad de datos: atributos, riesgos, indicador, métrica y meta	70
Tabla 14 Planeación estrategia de calidad de datos	71
Tabla 15 Planeación limpieza de datos	72
Tabla 16 Identificación de datos maestros y referenciales	74
Tabla 17 Definición de datos referenciales	74
Tabla 18 Planeación estrategia datos maestros y referenciales	75
Tabla 19 Identificación de metadatos	77
Tabla 20 Planeación estrategia de metadatos	78
Tabla 21 Arquitectura empresarial para la entidad afiliados	80
Tabla 22 Arquitectura de datos para la entidad afiliados	80
Tabla 23 Planeación estrategia de arquitectura de datos	81
Tabla 24 Seguridad de datos e información: identificación de activos	83
Tabla 25 Clasificación de activos por tipo y dimensión	84
Tabla 26 Planeación estrategia de seguridad de datos e información	85
Tabla 27 Definición de entidad, atributo y dominio	86

Tabla 28 Estrategia de modelado de datos	87
Tabla 29 Listado de indicadores medición de cultura organizacional.....	90
Tabla 30 Mapa de ruta para gobierno de datos e iniciativas	91
Tabla 31 Formato definición de indicadores.....	93
Tabla 32 Ejemplo definición y seguimiento de indicador	94

Abreviaturas y acrónimos

Abreviatura / acrónimo	Descripción
CMMI	Capability Maturity Model Integration, Modelo de Madurez de Capacidades Integrado
CCF	Caja de Compensación Familiar
CDO	Chief Data Officer
CD	Calidad de datos
CV	Cadena de valor
DAMA	Data Management Association, Asociación de gestión de datos
DANE	Departamento Nacional de Estadísticas
DGI	Data Governance Institute, Instituto de gobierno de dato
DMM	Data Management Maturity, Madurez en gestión de datos
DM	Datos maestros y referenciales
EPS	Entidad prestadora de salud
GD	Gobierno de datos
IBM	International Business Machines Corporation, Servicios y Consultoría de Tecnología Internacional
IPS	Institución prestadora de salud
OCDE	Organización para la Cooperación y el Desarrollo Económicos
SSF	Superintendencia de Subsidio Familiar
UGPP	Unidad de gestión de pensiones y Parafiscales
MINTIC	Ministerio de la tecnología y las comunicaciones
MD	Metadatos
TI	Tecnología de la información

1 INTRODUCCIÓN

La sociedad experimenta un momento de gran desarrollo tecno-científico, y en este mismo sentido lo hacen las empresas, las cuales viven un crecimiento exponencial (Ismael S., Malone M., 2016); y de igual forma los datos producidos por la humanidad ahora son mayores en frecuencia y volumen.

La habilidad para aplicar una correcta gestión de los datos se ha convertido en la nueva fuente de riqueza de la economía, de manera que ha surgido la necesidad de establecer una adecuada estrategia de gobierno de datos.

El activo más trascendental para la humanidad ahora son los datos, y por tal motivo se deben proteger, administrar, gestionar y desarrollar; es decir, establecer un gobierno que permita la adecuada gerencia de los mismos y lograr la correcta toma de decisiones como resultado final.

Si bien no existe una única manera y cierta de ejercer gobierno sobre los datos, se pueden establecer parámetros mínimos que permitan definir un camino de ruta hacia los objetivos definidos dentro del plan estratégico de la empresa.

Surgen de esta manera metodologías, que sirven como referencia para aplicar buenas prácticas en torno a la correcta gestión y administración del activo más importante para las organizaciones.

1.1 Justificación

En este texto, se plantea el valor de establecer un modelo estándar e integral de gobierno de datos, para aplicarse en cualquier tipo de empresa de un campo específico de la economía colombiana; abordando la temática desde un punto de vista conceptual y propositivo al uso de buenas prácticas de administrar los datos de forma correcta.

El marco de referencia de gobierno de datos pretende impactar de forma positiva la gestión empresarial, en la medida que se administren y controlen los datos efectivamente, las actividades de toma de decisiones serán acertadas ya que existirán los fundamentos y soportes necesarios.

El gobierno de datos es idóneo para desarrollar las capacidades empresariales y generar un ecosistema basado en datos, en la información y en el conocimiento; forjando de esta manera un valor definido hacia los clientes, los productos y los procesos.

Lo anterior, derivado de cambios estructurales en el entorno económico, empresarial y social. Al final, se realiza una serie de recomendaciones y conclusiones que impactan en la implementación del modelo definitivo.

1.2 Objetivos

1.2.1 General

Proponer un marco de referencia de gobierno de datos integral aplicado al sector de las Cajas de Compensación en Colombia, para sustentar las decisiones estratégicas.

1.2.2 Específicos

- Entender y justificar la necesidad de poseer un gobierno de datos.
- Analizar la situación actual del negocio para identificar las iniciativas clave y sus requerimientos de datos.
- Identificar y analizar los diferentes componentes que deberían ser parte de un modelo de gobierno de datos.
- Proponer un plan de implementación de corto, mediano y largo plazo para la conformación y mantenimiento del modelo de gobierno de datos.

1.3 Alcance y limitaciones

Tomando como punto de referencia una destacada empresa del sector de la seguridad social en Colombia, se construirá un marco de referencia para la ejecución de un programa de gobierno de datos, considerando cada una de las variables a tener en cuenta en este tipo de proceso.

Las recomendaciones se construyen con base a las exigencias generales de gobernar y controlar los datos producidos en las etapas de su ciclo de vida.

Debido a que el arte de gobernar es un proceso constante y continuo, los resultados tangibles no harán parte del alcance de este documento, ya que este tipo de decisiones son de largo plazo.

Los aspectos a incluir dentro del modelo son tan amplios que implican un desarrollo independiente, por tal motivo, componentes como calidad, seguridad y arquitectura se enuncian de forma estratégica y se detallan a un alto nivel, de tal manera que la implementación no hace parte como tal del presente estudio, sólo se abordan como un eslabón importante dentro del gobierno.

Se plantea el modelo operativo, la cadena de valor y ciclo de vida, desde un punto de vista netamente conceptual, incluyendo todos los elementos necesarios para un

proceso adecuado de implementación y desarrollo de gobierno de datos; soportados en las metodologías disponibles y proponer la aplicación práctica en la empresa seleccionada, que sirva de referencia para organizaciones similares.

1.4 Metodología

El documento empieza con la investigación conceptual y contextual de los datos, del gobierno de datos y algunos casos prácticos que permiten visualizar un panorama en esta materia.

Se seleccionó una empresa para aplicar la metodología, la cual es un actor relevante dentro del sistema de seguridad social de Colombia, contando con 62 años de creación posee un amplio ecosistema de servicios y programas que hacen factible y relevante la realización del gobierno de datos.

El proceso para generar la propuesta de un marco de referencia de un gobierno de datos, empieza con la identificación del estado de madurez de la empresa seleccionada y de esta manera generar el diagnóstico del estado presente para poder plantear un estado futuro.

Con el análisis de madurez es posible determinar el modelo operativo, los componentes, la cadena de valor y los demás elementos que contempla el gobierno de datos.

Se aplica la metodología de caso de uso, es decir, tomar como referencia una empresa del sector para proponer un modelo general que pueda ser usado por cualquier actor del sistema de la Compensación Familiar.

2 MARCO CONTEXTUAL Y CONCEPTUAL

2.1 Conceptos básicos

Existen muchas y variadas definiciones del tema expuesto en este documento, todas válidas y complementarias entre sí. Se tomarán los conceptos de la Asociación de Gestión de datos -DAMA¹- para definir “datos”, “gobierno de datos” y “gestión de datos”:

- Dato: representación de hechos, conceptos, instrucciones de manera formal y adecuada para la comunicación, interpretación y procesamiento (DAMA, 2017).
- Gobierno de datos: ejercicio de la autoridad y control -planeación, monitoreo y mejora- sobre la gestión de activos de datos (DAMA, 2017).
- Gestión de datos: desarrollo, ejecución, supervisión de planes, políticas, programas y prácticas que ofrece, controla, protege y mejora el valor de los datos durante todo su ciclo de vida (DAMA, 2017).

Sin desconocer que existen otras definiciones de personas e instituciones igualmente válidas.

Los siguientes términos son usados durante la trayectoria del documento:

- Entidad: aquello que es importante para la organización, elementos cuya recolección es de valor estratégico.
- Atributo: características que son parte de la esencia de las entidades.
- Dominio: medio para estandarizar las características de los atributos.
- Custodio: desde el punto de vista técnico, responsable de administrar el dato.
- Propietario: desde el punto de vista de negocio, responsable de recolectar o crear el dato.

2.2 Sobre los datos

Si bien es cierto que los datos siempre han existido en toda la historia de la humanidad, la forma en la que se procesan y se tratan es diferente en cada época, las capacidades tecnológicas han permitido maniobrar un volumen alto de información; y encontrar nuevas formas de gestionar los miles y miles de datos que se producen en todo momento.

En Mayer-Schönberger (2013), se menciona que el registro contable más antiguo se encontró hace 5.000 años AC, donde los mercaderes sumerios utilizaban pequeñas cuentas para representar bienes en venta; situación que nos permite

¹ En inglés: Data Management Association

concluir que los datos siempre han estado presentes en los procesos de la evolución de la humanidad.

La sociedad nunca había experimentado que la tecnología avanzara a un ritmo tan acelerado como el actual. El crecimiento y desarrollo tecnológico es cada vez más exponencial y ya es común hablar de inteligencia artificial, aprendizaje automático, algoritmos, internet de las cosas, *Big Data*, etc., de acuerdo con el análisis realizado por Salim (2016).

El profesor Harari (2014) afirma que en los últimos 500 años se ha vivido la “revolución científica”, en la que la humanidad empieza a desarrollar un poder inimaginable sobre el entorno, incluso deja de lado la selección natural para pasar a lo que llama un “diseño inteligente”. Es decir, que en los últimos años la tecnología se ha desarrollado más que en varios siglos de historia.

En este contexto de gran avance tecno-científico, el 90% de los datos generados por la humanidad se han producido en los últimos 5 años. Existen numerosos sensores ubicados en dispositivos fijos y móviles: computadores, máquinas, vehículos, celulares, etc., los cuales recopilan, almacenan, envían y gestionan datos de cada operación que realizan los usuarios.

Incluso autores como Harari (2016) hablan de una nueva religión, “la de los datos”, porque se creará más en ellos que en las personas, se habla más de libertad de información que de los mismos seres humanos; por lo tanto, la economía del futuro se basará en los datos y en el conocimiento que se pueda derivar de su uso.

La economía centrada en datos que se ha estado gestando desde hace unos años, se convertirá en una nueva fuente de empleos, industrias y avances que cambiarán el desempeño social.

Las grandes empresas multinacionales que hoy en día dominan la economía, ya no son aquellas que se basan en la propiedad, son las que se soportan en los datos, ejemplos como Google, Amazon, Uber, Airbnb, Rappi, Alibaba, Facebook, etc.; son modelos que dan cuenta de esta afirmación.

En Harari (2016), se introduce el concepto de “datoísmo” o “dataísmo”, término acuñado para describir que el futuro de las organizaciones son los datos y las nuevas reglas de la humanidad van a estar determinadas por aquellos que logren explotar eficientemente los datos que se producen.

Los datos se convirtieron en la fuerza que está moldeando el entorno empresarial, político, social, educativo y tecnológico. Las grandes ideas y las mejores decisiones se toman por aquellos cuya fuente de inspiración son los datos.

De acuerdo con Soares (2011) las decisiones basadas en los datos organizacionales, tendrán un mejor desempeño en las ventas. El factor diferencial de los negocios ya no es la tierra, la mano de obra o el capital en sí; se cita a los datos como el recurso clave y el factor que puede llevar a la distinción y continuidad en los negocios.

En el planteamiento realizado por Muñoz (2015); describe la necesidad de definir un modelo organizacional basado en datos, para que las disciplinas y capacidades se articulen como un todo, y que los procesos sean eficientes en torno a la administración de los datos, desde la producción hasta el consumo de ellos.

El profesor Mayer-Schönberger (2017) introduce el concepto de capitalismo de los datos, en el cual plantea que los datos han sustituido en gran medida a los precios como indicador de la economía y que los mercados ricos en datos dejarán obsoleta a la empresa tradicional.

La evolución de los datos se puede entender como inherente a los procesos sociales y empresariales, la diferencia con la sociedad tecnológica actual es la capacidad de uso y desarrollo que se les pueda dar; incluyendo todas aquellas metodologías que nacen en torno a su custodia, gestión y administración. Tanto la continuidad de las empresas como la aparición de nuevos negocios, están atados a la correcta administración y gestión que se realice sobre los datos.

Destacando el valor de los datos, el Estado Colombiano por ejemplo estableció el Conpes 3920 en la que establece la “política nacional de explotación de datos” para el sector público, con el objetivo de aumentar el aprovechamiento de datos; lo que lleva a todos los actores a pensar en una dinámica de masificación y cultura de los datos.

Así las cosas, los datos se convierten en un elemento fundamental del proceso cultural y empresarial de la humanidad, que tienen diferentes matices para lograr su adecuada gestión y administración.

2.3 Sobre el gobierno de datos

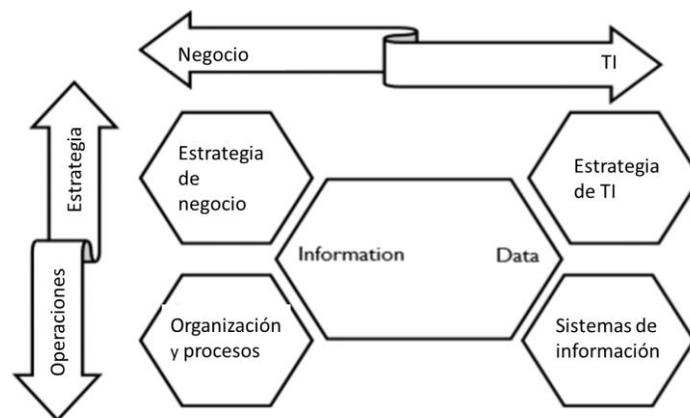
Ahora bien, entendido el nuevo orden mundial donde los datos serán los nuevos reguladores de la oferta y la demanda, es necesario preguntarse ¿cuáles capacidades se desarrollan para administrar estos desafíos?, ¿cuáles son los roles y responsabilidades a incluir?, ¿existen políticas de gestión de datos? ¿cuál es el modelo operativo?, ¿cuál es la estrategia cultural para la toma de decisiones a partir de información y conocimiento? y ¿cómo los procesos entran en etapas de reingeniería para garantizar que las organizaciones se articulen en una economía que, como se ha mencionado, está sustentada en los datos?

En este punto se podría llegar a la conclusión que los cambios forjados por los avances tecno-científicos, han desencadenado un crecimiento exponencial en la cantidad y volumen de datos; y que dicha situación lleva a plantearse nuevos retos para implementar un esquema de gobernanza de datos.

Con la evolución de las capacidades, y la respuesta tentativa a las preguntas anteriores, es posible dirigirse hacia un interrogante y reto aún más grande: ¿cuál es el modelo estándar e integral que garantice una correcta administración y gestión de los datos?

De esta manera se han venido estructurando una serie de metodologías, que no se entienden como recetas únicas, sino complementarias, resulta importante desarrollar la capacidad de adaptar el referente elegido a la industria específica por las condiciones económicas, sociales y organizacionales de las empresas, incluso tomar elementos de cada modelo para constituir uno que resulte propio y único, alienado a las exigencias y los objetivos definidos; entre otros modelos tenemos:

- Modelo de alineación estratégica: incluye actividades para involucrar e integrar al negocio con la capacidad de tecnología, tanto a nivel operativo como estratégico; se destaca a los datos e información como elementos centrales y protagonistas del modelo; permitiendo de esta manera una completa alineación de los procesos, tal como se muestra en la ilustración 1.



*Ilustración 1 Modelo alineación estratégica
Fuente: (Henderson y Venkatram, 1999)*

- Modelo de la información de Amsterdam: de acuerdo a la ilustración 2, se ubica al negocio junto con el área de TI para desarrollar las capacidades de gobierno de información, se destacan procesos a nivel estratégico, táctico y operativo. Nótese que en la parte inferior estaciona a la calidad de datos como asiento de

todo el modelo. Llevando también el componente de arquitectura como un elemento visible dentro del planteamiento.



*Ilustración 2 Modelo de la información de Amsterdam
Fuente: (Abcouwer, Maes & Truijens, 1997)*

- Pirámide de Peter Aiken: como se puede apreciar en la ilustración 3, el modelo se basa en fases, en las cuales va estructurando su modelo de gobierno de acuerdo a sus capacidades técnicas, de procesos y de personas; es decir, una evolución de acuerdo a la disponibilidad de recursos de la empresa.

Fase I: modelo inicial con actividades básicas del gobierno (almacenamiento, seguridad, integración de datos)

Fase II: identificación de nuevos elementos fundamentales como la calidad y metadatos.

Fase III: para alcanzar una evolución de la fase II se requiere involucrar la iniciativa del gobierno como tal (gestión de contenidos, datos maestros y de referencia)

Fase IV: una vez establecidas las capacidades iniciales, se sumerge en la necesidad de implementar prácticas avanzadas que permitan el desarrollo de los datos y su exploración y explotación (Big Data, Análítica predictiva y prescriptiva, Machine Learning). Tal como el autor lo denomina “pirámide de oro de la gestión de datos”

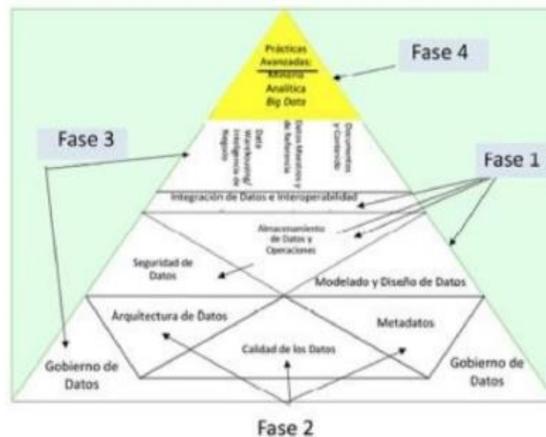


Ilustración 3 Pirámide de Peter Aiken

- DAMA (Data Management Association): es una asociación sin fines de lucro que promueve las prácticas de gestión de datos a nivel mundial. Las actividades que se listan dentro de la metodología, incluyen: definir el GD para la organización, valoración del estado actual, realizar alineación y descubrimiento de negocio, descubrir los datos críticos, desarrollar la estrategia, definir el modelo de operación, enunciar políticas y objetivos, establecer los proyectos de datos.

Durante este proceso se incluye el glosario de negocio, así como las actividades de gestión de cambio y la identificación de los patrocinadores.

- IBM (International Business Machines): esta metodología presenta una serie de pasos para la implementación del GD, tal como se muestra en la ilustración 4:

Definir el caso de negocio, obtener el respaldo de los responsables, realizar el modelo de madurez, construir un mapa de ruta, estructurar el diccionario de datos, entender los datos, crear el repositorio de metadatos y la definir las métricas. Establece el gobierno de datos maestros e incluye la identificación de responsables, gestión de calidad y datos maestros como tal. Por último, el componente de analítica, seguridad, ciclo de vida de la información y la medición de resultados.

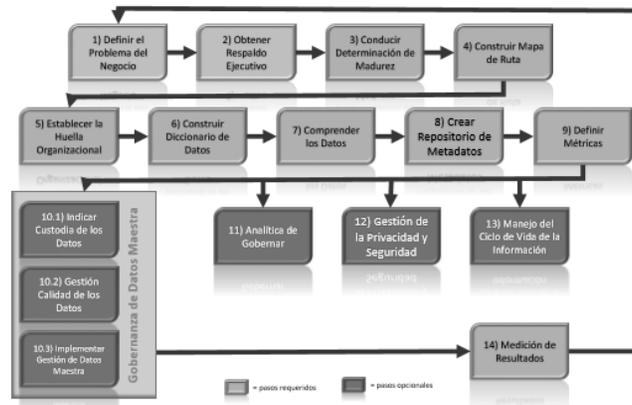


Ilustración 4 Propuesta gobierno de datos IBM
Fuente: (IBM, 2010)

- DGI (Data Governance Institute): para el Instituto de Gobierno de datos se incluyen 10 componentes necesarios para alcanzar la correcta aplicación de la metodología, como se puede apreciar en la ilustración 5:

Misión y visión: hacia donde se dirige el programa, por qué se está implementando la iniciativa.

Objetivos, métricas y estrategia: mediciones que permitan apalancar el seguimiento del GD y cómo garantizar que los esfuerzos realizados estén teniendo el impacto en los procesos.

Reglas de datos y definiciones: políticas, estándares, reglas de negocio, requerimientos de datos.

Flujo de decisiones: proceso de decisiones cruciales (facilitar y recopilar).

Responsables: cada una de las actividades relacionadas, existen roles y responsabilidades asignadas a un cargo o persona específica.

Controles: estándares definidos en las medidas de control para asegurar el ciclo de vida de los datos, así como las decisiones efectuadas.

Partes interesadas: identificación de todos los interesados en el flujo de los datos e información, desde la creación hasta el uso mismo.

Oficina de gobierno de datos: aquella entidad encargada de orquestar, alinear y articular las decisiones, definiciones y monitoreo del programa de DG.

Dueños del dato: aquellas entidades que son propietarios y por lo tanto responsables de los datos, los cuales se integran con los niveles superiores para el cumplimiento de las políticas.

Procesos de gobierno de datos proactivos, reactivos y continuos: ejecución de un programa estandarizado, documentado y repetible.

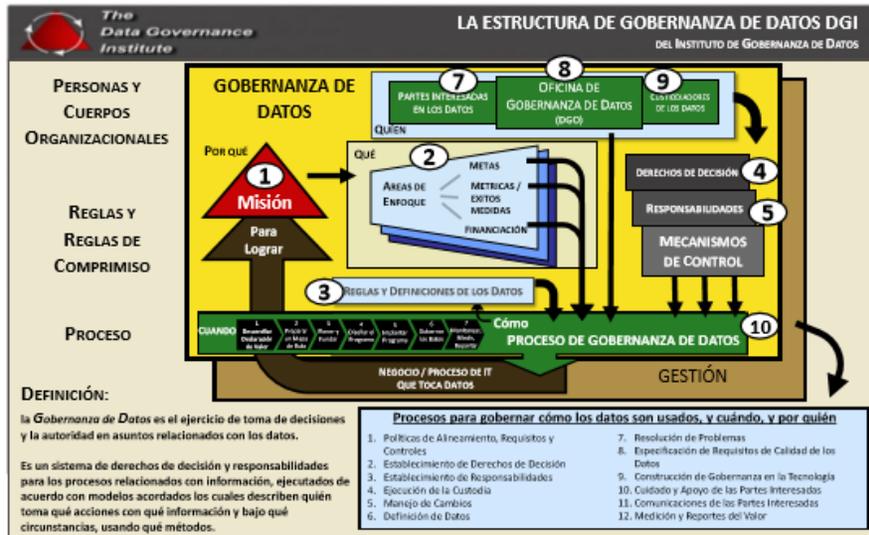


Ilustración 5 Estructura de gobierno de datos DGI
Fuente: (DGI, 2014)

Realizado este recorrido académico sobre las metodologías indagadas, se podría concluir que el modelo de gobernanza de datos integra una serie de disciplinas y áreas de conocimiento que se involucran en un mismo horizonte de tiempo para lograr los objetivos estratégicos de la organización.

El gobierno abarca cambios orgánicos, los cuales contienen nuevas responsabilidades acompañados de cargos que en otra época no existían, por ejemplo, científico, custodio y vigilante de información; son roles específicos que conocen el negocio y los flujos de datos para traducirlos en un lenguaje sencillo y accesible a los tomadores de decisiones. Soares (2015), analiza el nacimiento de los nuevos cargos que han ido apareciendo por la oportunidad de administrar una organización basada en datos.

Vale la pena resaltar que el modelo de gobernanza de datos, es una herramienta o un habilitador que permitirá realizar la adecuada gestión de datos, es decir, el hecho de ejecutar dicho programa no garantiza que se solucionarán todos los problemas relacionados con los datos. Se constituye en un marco de referencia para gestionar de una forma más adecuada.

2.4 Sobre la gestión de datos

Las actividades en torno a los datos se han convertido en una nueva función dentro de la cadena de valor empresarial.

En las definiciones de Gartner² se precisa que la gestión de datos consiste en todas las prácticas, técnicas y herramientas que permiten el acceso y entrega de datos en todas las áreas de la organización y así cumplir con todos los requisitos de los clientes finales.

La estructura propuesta por la DAMA (2017) incluye una serie de áreas de conocimiento que articuladas entre sí logran una gestión integral de datos (arquitectura, modelamiento, almacenamiento, seguridad, integración e interoperabilidad, documentos y contenidos, datos maestros y de referencia, inteligencia de negocios, metadatos y calidad); de acuerdo a la ilustración 6.



Ilustración 6 Áreas de conocimiento
Fuente: (DAMA, 2017)

La metodología también aborda un diagrama de contexto que permite a la organización la visualización de las macro-actividades para la implementación de las iniciativas, empezando por las entradas, las actividades, los entregables, definiendo proveedores, participantes y clientes; soportados en técnicas, herramientas y métricas (ver ilustración 7).

² Recuperado de: <https://cutt.ly/pf5h7az>

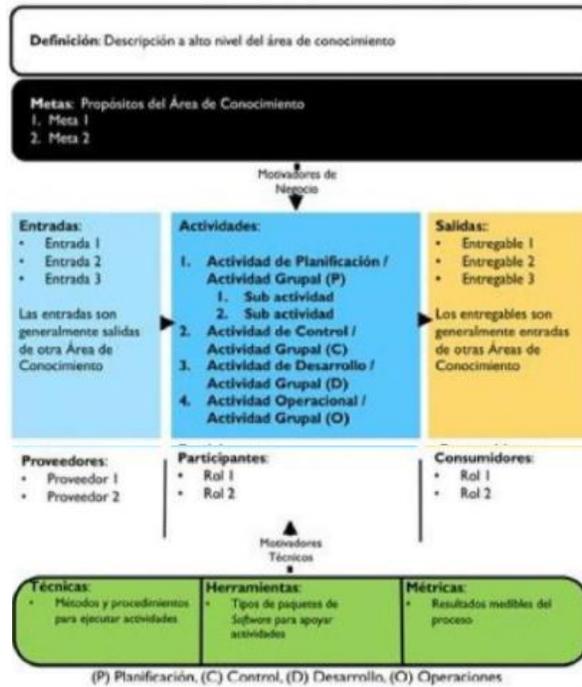


Ilustración 7 Diagrama de contexto
Fuente: (DAMA, 2017)

E incluye un hexágono de factores ambientales, los cuales se convierten en aquellos inductores que garantizan el éxito de la implementación de una decisión de este tipo, ya que se requiere articular los procesos, con la tecnología y las personas para alcanzar las metas planteadas (ver ilustración 8).



Ilustración 8 Factores ambientales
Fuente: (DAMA, 2017)

2.5 Sobre las Cajas de Compensación Familiar

Los orígenes de este tipo de empresa se remontan a 1954 por medio de una iniciativa privada del Sindicato Antioqueño para aliviar las cargas económicas de los

trabajadores; se estableció el subsidio monetario y la prestación de servicios sociales para mejorar la calidad de vida.

Las Cajas de Compensación son Corporaciones privadas, adscritas al sistema de seguridad social en Colombia, vigiladas por la Superintendencia del Subsidio Familiar; fuertemente reguladas a través de un robusto marco normativo (Decreto Único Reglamentario 1072 de 2015, 1077 de 2015, entre muchos otros).

Los ingresos provienen del recaudo del 4% (recaudo parafiscal³ “contribución proporcional al ingreso” que toda empresa formal paga sobre el pago realizado a cada trabajador) y la prestación de servicios (propios o en alianza con otras entidades públicas y privadas). Por tratarse de recursos parafiscales, las convierte en corporaciones privadas administradoras de recursos públicos.

Son aliadas estratégicas con el gobierno nacional y con las empresas aportantes, ya que se desarrollan programas en torno al beneficio de los trabajadores afiliados y sus personas a cargo, creando portafolios integrales que se adecuen a las necesidades particulares de cada agente involucrado en el proceso del desarrollo humano.

La posibilidad de prestación de servicios es amplia y la estrategia se ejecuta de acuerdo a las decisiones administrativas de cada Consejo Directivo, el portafolio de las CCF cubre las exigencias del ser humano a lo largo de su ciclo de vida; la diversidad de operaciones se encuentra concentrada en:

- Otorgamiento de subsidios: cuota monetaria⁴, en especie⁵, en servicios⁶.
- Servicios sociales: educación, recreación, crédito, turismo, bibliotecas.
- Servicios auto-sostenibles: mercadeo (supermercados y medicamentos), salud (IPS y EPS).
- Participación accionaria en otras empresas de la seguridad social como Pensiones y Cesantías, Salud; y otras como el sector de Educación, del Microcrédito, Recaudo de Parafiscales, etc.
- Administración de fondos sociales: Fondo para la atención integral a la niñez y jornada escolar complementaria, Fondo de Solidaridad de Fomento al Empleo y Protección al Cesante, Fondo de Solidaridad y Garantía del Sistema General de Seguridad, Fondo para el Subsidio Familiar de Vivienda de Interés Social.

³ <https://www.corteconstitucional.gov.co/relatoria/2009/c-307-09.htm>

⁴ Valor pagado en dinero por cada persona a cargo menor de edad del trabajador afiliado, modificado anualmente de acuerdo a la normatividad vigente.

⁵ Útiles escolares, descuentos en productos de la canasta básica.

⁶ Becas en pesiones, matrículas, entradas a parques recreativos y hoteles.

- Administración de recursos de entidades externas: recursos públicos y privados para ejercer programas de desarrollo y responsabilidad social, dirigidos a población focalizada o abierto al público en general.

En la siguiente ilustración se muestra el diagrama Canvas⁷, reflejando el modelo operativo de una organización de este tipo. El valor prometido es mejorar la calidad de vida de afiliados en primera instancia y los no afiliados en un segundo momento; a través de medios de contacto tanto físicos como virtuales se ofrecen los productos y servicios enmarcados dentro del marco normativo, destacando a su vez el trabajo articulado con las empresas, el gobierno y la sociedad para el cumplimiento de la razón de ser de una CCF.



Ilustración 9 Diagrama Canvas para una CCF
Fuente: elaboración propia

Actualmente existen 43 Corporaciones de esta naturaleza distribuidas en todo en el territorio nacional, de acuerdo a la regionalidad⁸ consagrada en la normatividad vigente. Presentan variantes de acuerdo a la cantidad de servicios ofrecidos y a la capacidad financiera.

El caso de uso se está aplicando sobre una organización cuyo impacto social es relevante, con una estructura bastante compleja al tener diversidad de servicios y por lo tanto sistemas transaccionales diversos.

⁷ Documentación del modelo de negocio existente para entender a manera general los principales elementos de la empresa y sus relaciones.

⁸ El concepto de regionalidad aplica únicamente para la afiliación de las empresas y trabajadores; la prestación de servicios se puede dar en cualquier lugar de la geografía nacional.

Se generan más de 60 mm de datos en la operación mensual y se están implementando las herramientas que permitan manejar tal magnitud de datos, los cuales presentan una tasa de crecimiento de 10%⁹.

Como ya es evidente las CCF, juegan un papel primordial dentro del sistema de la seguridad social del país, razón por la cual la adecuada gestión de los datos resulta relevante. El conocimiento de las necesidades de los afiliados lo tienen las Cajas de Compensación, ya que, desde la afiliación hasta el uso de los servicios, se están recolectando datos.

A diciembre de 2019, las 43 corporaciones presentaban las siguientes cifras consolidadas:

- 685.895 empresas afiliadas
- 9.670.065 trabajadores afiliados
- 11.464.328 personas a cargo afiliados
- \$20,7 billones de ingresos operacionales y por ganancias
- \$20,1 billones de costos y gastos
- \$0,6 billones por remanente del ejercicio
- \$23.5 billones en activos
- \$12,2 billones en pasivos
- \$11.4 billones en patrimonio

2.6 Del estado del arte

Ante los requisitos nacientes y crecientes de gestión de datos, las CCF han venido estableciendo una serie de medidas que les permita administrar de forma correcta los datos que se generan dentro del ciclo de vida; determinaciones que se aplican desde el interior para cumplir con los requisitos de ley y de conocimiento del cliente individual y empresarial.

En este contexto, autoridades a nivel extra corporativo se han dado a la tarea de ejercer recomendaciones y directrices que permitan desarrollar la gestión, control y gobierno de los datos dentro de estas organizaciones.

Las fuerzas que regulan las actividades del GD se puede denominar un tipo de “gobernanza extra corporativa” (Bernal, 2019) y hace referencia precisamente a aquellas acciones que ejercen un control sobre las Corporaciones, aplicando lineamientos, reglas, definiciones comunes, estándares, etc.; indicaciones

⁹ Dato estimado sobre el volumen de crecimiento de coberturas y transacciones en Salud, Educación, Recreación, Crédito, Supermercados y Medicamentos.

puntuales que se aplican para homologar términos y conceptos comunes y facilitar de alguna manera las tareas de vigilancia y control.

Entre otras tenemos:

- OCDE (Organización para la Cooperación y el Desarrollo Económicos): propone los principios de gobierno corporativo, cuyo objetivo es ayudar a evaluar y mejorar el marco legislativo, reglamentario e institucional. Este tipo de propuestas se convierten en un referente dentro de las acciones de empresas que deciden adoptarlas, y el gobierno de datos no es ajeno a incorporar buenas prácticas que ayuden a la adopción y seguimiento de los programas y proyectos.
- Ministerio de las Tecnologías de la Información y la Comunicación, MINTIC: emite la guía técnica que contiene y da un contexto del GD; si bien define los aspectos que las instituciones del sector público han de tener en cuenta para un adecuado GD, es aplicable a cualquier tipo de entidad interesada en el tema.

Define los ámbitos del gobierno del dato visualizados en la ilustración 10, se destaca la gobernanza, calidad, migración, ciclo de vida y administración de los datos:



*Ilustración 10 Ámbitos del gobierno del dato
Fuente: (Mintic, 2014)*

- Superintendencia del Subsidio Familiar, SSF: se han realizado varias iniciativas de generar directrices básicas que permitan ejercer un gobierno de datos para el conjunto de las 43 Corporaciones, por citar algunas:

Emisión de circulares normativas¹⁰ en torno a los reportes de información, con un lenguaje común y unificado.

Manuales de diligenciamiento de información, aplicando estándares de recolección de información.

¹⁰ Actualmente se encuentra en vigencia la Circular Externa No. 0007 de 2019 versión 2

Identificación, definición y comunicación de términos comunes y transversales a las empresas del sistema, es decir, un glosario de negocio único para el sector¹¹.

Cabe anotar que la SSF, hace parte del sistema estadístico nacional del DANE, lo cual implica la aplicación de estándares internacionales en torno al tratamiento de los datos, así como los mecanismos de recolección. Estas estrategias unificadas permitan tener estadísticas de valor para todo el sistema, lo cual se logra con estos proyectos.

- Unidad de Gestión Pensional y Parafiscales: la UGPP realiza actividades de inspección y vigilancia, dando indicaciones y directrices en torno a la gestión de la información y elaboración de reportes referente al recaudo de aportes del 4%.

Al igual que la SSF, esta entidad también determina estándares que permitan tener un lenguaje unificado con conceptos uniformes en torno a los datos, facilitando su labor de vigilancia, pero a su vez ejerciendo parámetros de gobierno sobre las CCF.

A nivel empresarial tanto en el sector público como privado, se han identificado e investigado los siguientes casos, sin decir que sean los únicos:

- Hospital general de Medellín (Vallejos, & Soto, 2019): en este caso los autores proponen un diseño del programa de gobierno de datos, en el cual se identifica para esta institución los objetivos, la metodología, roles y responsabilidades, procesos, políticas, inventario de datos, diccionario e indicadores clave de desempeño.
- En Portilla (2016), se ha realizado un estudio para analizar los efectos de la integración de un programa de gobierno de datos, con el sistema de gestión de calidad en Positiva Compañía de seguros.

Si bien es cierto que este tipo de servicios no existen en todas las CCF, es importante el desarrollo y conclusión de casos que miden el impacto del GD dentro de la organización.

- El Instituto Distrital de Gestión de Riesgos y Cambio Climático¹², propone una política de GD, en la cual genera las herramientas para verificar, controlar y mejorar la calidad; estableciendo mecanismos para lograr que los datos eleven

¹¹ <https://www.ssf.gov.co/transparencia/informacion-de-interes/glosario>

¹² Recuperado de: <https://cutt.ly/Jf5hvHn>

su calidad. Define áreas a las que tiene alcance, así como un período de tiempo de implementación.

- La gobernabilidad de datos como apoyo en la gestión de datos de instituciones de educación superior, es una propuesta de Osorio, Guerrero, González-Zabala (2017); en la cual plantean un marco de trabajo para empresas de este sector, ya que requieren de datos para definir y evaluar sus estrategias, tiene un enfoque hacia la calidad para soportar las actividades de toma de decisiones.

Se destacan tres etapas:

Definición de enfoque (políticas, estándares y estrategias, calidad de datos, privacidad, cumplimiento y seguridad, arquitectura e integración de datos, bodega de datos e inteligencia de negocios).

Selección de dimensiones de calidad (agrupación de los factores de calidad).

Validación del plan (análisis del entorno, evaluación de la gobernabilidad y calidad de los datos, diseño del plan y validación).

De ahí surge la exigencia de estructurar un modelo o marco de referencia de gobierno de datos, que sirva para gestionar y desarrollar los datos y contribuir de esta manera al cumplimiento de los objetivos estratégicos de las CCF; desarrollando productos y servicios que se ajusten al bienestar de millones de personas afiliadas en Colombia, mejorar la productividad y garantizar el cumplimiento normativo.

3 DIAGNÓSTICO

El capítulo del diagnóstico se incluye el análisis de la situación actual para identificar las oportunidades e insumos del modelo de gobierno; el objetivo es determinar el camino por dónde empezar porque son muchas las tareas que se deben accionar.

3.1 Modelo de madurez

En Ampuero (2018) la invitación antes de empezar a construir o definir un modelo de gobierno es estructurar un análisis interno, identificar cuál es el estado de madurez de la empresa y conocer los aspectos por mejorar, reformar, redefinir y crear.

Una vez se conozca dicho punto de partida, será posible combinar las capacidades y disciplinas que formarán parte del modelo de gobierno de datos. La identificación del nivel de madurez se convierte en el primer paso por desarrollar dentro del análisis de construcción de la propuesta de gobierno de datos. No es posible pensar en aplicar cambios si no existe un conocimiento detallado de cómo la empresa se encuentra actualmente.

Con el propósito de analizar el estado actual se toma como referente la metodología del CMMI, específicamente el DMM para realizar la valoración de percepción. Sin desconocer que existen otros métodos igualmente válidos (DAMA, IBM, Stanford, etc.), los cuales tienen el mismo objetivo y es lograr medir el punto de partida para la implementación del GD.

Los niveles de madurez presentados por la metodología del CMMI son:

Nivel 1. Realizado: los procesos se realizan sin planificación, se aplican a todas las áreas de la organización, pero son reactivos, no existe un enfoque estándar y común para administrar o compartir datos.

Nivel 2. Gestionado: las técnicas se planifican y ejecutan de acuerdo a políticas, se monitorean, controlan y se usan en toda la organización. Hay apariencia de gobernanza de datos y una preocupación por la calidad de los datos.

Nivel 3. Definido: los procesos estandarizados se siguen constantemente y se tienen métodos específicos de acuerdo a las pautas de la organización. Se ha definido un modelo de gobierno donde las reglas de negocio son capturadas por un responsable de gestión de datos.

Nivel 4. Medido: las métricas del proceso se definen y se aplican, se ha determinado e implementado un modelo de gobierno con un área responsable y enfocada a la calidad de los datos.

Nivel 5. Optimizado: el modelo de gobierno se optimiza, se identifican mejoras y se comparten las mejores prácticas. Existen objetivos cuantitativos para mejorar los procesos y se revisan continuamente. La organización es segura, ágil e inteligente y la información se comparte con total transparencia.

Al aplicar la metodología del nivel de madurez del CMMI se busca alcanzar dos objetivos:

- Determinar el nivel actual de la organización para la implementación del GD.
- Identificar inductores y temáticas a priorizar dentro del modelo de GD.

3.2 Desarrollo del caso práctico

Se aplicó un formulario tipo encuesta de análisis exploratorio con el fin de identificar la percepción del gobierno de datos en la organización; usando la escala de Likert¹³.

Para el caso de estudio, se adaptaron 114 preguntas en sentido afirmativo, indagando los componentes de estrategia de gestión de datos, comunicación, función de la gestión de datos, gestión de gobierno, glosario o diccionario, gestión de metadatos, estrategia de la calidad, perfiles, evaluación de la calidad, limpieza o depuración, operaciones, plataforma y arquitectura, integración, históricos y procesos de soporte.

Por cada uno de los públicos objetivos identificados se presenta la responsabilidad dentro de la organización, el resultado frente a la percepción y la identificación de oportunidades para alimentar el modelo de gobierno, actividades futuras y brechas que se abordan para apalancar la estrategia de gestión de datos.

El público objetivo se encuentra separado de acuerdo al rol que desempeñan dentro de la organización: transversal estratégico, transversal operativo, tecnológico, negocio y operativo.

1. Transversal estratégico

Responsabilidad: asesorar el direccionamiento estratégico en cada una de las áreas, estudio de tendencias y mejores prácticas, definición de indicadores y metas a los objetivos planteados.

Cargos: especialista de estrategia, especialista de proyectos y especialista de investigaciones sociales.

¹³ Método de medición utilizado con el fin de evaluar la opinión y actitudes de las personas. Tiene su origen por el psicólogo Rensis Likert.

2. Transversal operativo

Responsabilidad: administrar y coordinar la información interna y externa de la empresa, así como realizar estudios que apalancan y soportan las decisiones estratégicas.

Cargos: analista de información, analista de modelos y auxiliar de información.

3. Tecnológico

Responsabilidad: gestionar la implementación de las iniciativas y mejoras tecnológicas de cada una de las áreas de la empresa, incluyendo la bodega de datos.

Cargos: coordinador de la oficina de proyectos, analistas de proyectos, especialistas de extracción de datos.

4. Negocio

Responsabilidad: gestionar la producción estadística en cada una de las áreas asignadas, así como la supervisión de los sistemas transaccionales en los que se realiza la captura de datos.

Cargos: jefes financieros, analistas de costos, coordinador de información, analista de información, auxiliar de información.

5. Operativo

Responsabilidad: producir los datos asociados en los formatos asignados.

Cargos: jefe de eventos, auxiliar de información, analista de información.

3.3 Definición del nivel de madurez

Como resultado de este ejercicio y como se evidencia en la ilustración 11, se puede concluir que la organización se encuentra en el **-nivel 2. Gestionado-**, porque si bien existen métodos en torno al manejo y la gestión de los datos, aún no existe el programa de gobierno establecido como tal, se han realizado actividades de depuración y calidad de datos, pero aún no forma parte de un proceso formal. Aún no se han desplegado todos los componentes de un gobierno como tal.



Ilustración 11 Valoración del nivel de madurez
Fuente: Elaboración propia

De esta manera se da cumplimiento al primer objetivo de la aplicación de la metodología, al determinar un nivel de madurez inicial en torno a la gestión de los datos; si bien no existe un proceso formalizado de GD, se desarrollan actividades inherentes y propias del GD.

Los resultados cualitativos y de percepción por cada nivel evaluado, son:

1. Transversal estratégico

Se percibe un alto apoyo de la Dirección a la estrategia de gestión de datos, pero aún el direccionamiento estratégico no está alineado, lo cual se debe en gran medida por la falta de entendimiento entre las partes interesadas y las pocas capacidades físicas y de personal para implementar la estrategia.

2. Transversal operativo

No se percibe condiciones para la ejecución de un programa de gestión de datos, no existe entendimiento ni acuerdo entre las partes interesadas. Los roles y responsabilidades aún no se encuentran maduros dentro de los dominios de datos.

3. Tecnológico

Se destaca la percepción de no existencia de glosario de datos y mapeo de términos. Ausencia de modelos de datos y estrategia de metadatos. Existe el área de gestión de la seguridad de la información, pero aún no está lo suficientemente alineada a la estrategia de gestión de datos. Aún no se posee un proceso formal de solicitud de requerimientos relacionados con los datos.

4. Negocio

Existen procesos de captura, evaluación y verificación de datos con requisitos previos; aunque no se perciben definiciones de datos estándar, ni funciones y responsabilidades claras, ni tampoco métodos de modelado.

5. Operativo

La percepción general es que no existe comunicación formal de la gestión de datos en la Corporación.

3.4 Oportunidades e inductores

Al aplicar la metodología y al ser preguntas de tipo exploratorio, se pueden identificar las siguientes oportunidades, las cuales se convierten en insumo para realizar la priorización de los entregables del marco de referencia de GD:

1. Transversal estratégico

- Establecer un programa de GD con las condiciones mínimas para su implementación, incluyendo el flujo de poder de decisión.
- Constituir un comité de GD con las partes interesadas, identificando un líder único y visible, líderes por cada uno de los frentes de trabajo y la definición de los componentes a incluir.
- Desplegar una estrategia de capacitación a un público objetivo para generar el entrenamiento que permita la madurez del uso, gestión y administración de los datos.
- Identificar mecanismos que permitan la colaboración de las partes interesadas.
- Definir macro-procesos dentro de una estrategia definida de la gestión de datos.
- Consolidar un proceso formal de auditoría y calidad, que permita validar y asegurar los datos en todo el ciclo de vida.
- Ejecutar esquemas de gestión de riesgo asociado a los datos, sobre todo en aquellos mapeados como críticos.

2. Transversal operativo

- Definir roles y responsabilidades que permitan establecer las funciones, procesos y aportes de cada área interesada.
- Liderar la estrategia de mapeo de términos de negocio y de referencias cruzadas para lograr alineación única de los datos. Así como el acompañamiento corporativo a la definición de metadatos, glosario de negocio y diccionario de datos.
- Aplicar método estándar de perfilado de datos que permita realizar un adecuado estudio, análisis y creación de resúmenes. Así como la generación de estrategias de capacitación en torno al perfil de los datos.

- Desarrollar estudios e indicadores para medir el impacto (en costo y reputación) de los problemas de calidad de datos. Igualmente, la ejecución de la política de limpieza de datos ejercida desde el comité de gobierno.
- Apoyar estrategia de integración y consolidación de los datos.

3. Tecnológico

- Definir la gestión de la demanda, proceso de solución ante los nuevos requerimientos de datos e información.
- Generar política y estrategia de metadatos, diccionario, glosario y modelado de datos, desde el punto de vista tecnológico y de negocio, incluyendo sistemas transaccionales. El cual será de referencia para todo proyecto relacionado con datos.
- Fortalecer el programa de Seguridad de la Información existente, orientado hacia la identificación de los activos de datos e información.

4. Negocio

- Desarrollar una estrategia para identificar el ciclo de vida de los datos desde su producción hasta su consumo y posterior archivo.
- Implementar indicadores de gestión para el análisis de las metas del gobierno de datos.
- Generar las necesidades de establecer puntos de control en la captura de los datos de los sistemas transaccionales.

5. Operativo

- Desarrollar una estrategia de comunicación para permear a un público objetivo con las nuevas tendencias y ejercer un adecuado gobierno de los datos, incluyendo lo concerniente a calidad, seguridad y arquitectura.
- Establecer planes y políticas que permitan evaluar los datos, definir márgenes de tolerancia y análisis de causas fundamentales; articular a todos los autores involucrados.
- Implementar una política y herramientas tecnológicas de consolidación de datos, así como de datos históricos.

4 MARCO DE REFERENCIA

Una vez identificado el punto de partida con el nivel de madurez de la empresa analizada, se procede a visionar el estado futuro, es decir, a proponer un esquema de operación inicial, el cual se ajustará a los propósitos específicos y recursos de cada caso.

Dentro del estado futuro se propone el establecimiento de elementos, componentes y aspectos a desarrollar que, integrados entre sí orquestan la estrategia del gobierno de datos, desde un punto de vista netamente estratégico.

El gobierno se entiende, alinea y comunica desde la estrategia de gestión de datos de la Corporación¹⁴, lo cual es fundamental para diagnosticar las capacidades y poder ejecutar las estrategias derivadas en torno a los datos; no puede actuar independiente, requiere del compromiso y liderazgo a todo nivel.

El camino que se propone seguir para la implementación de un modelo de GD, es el siguiente:

1. Planeación:

- a. Determinar factores de éxito.
- b. Estructurar el modelo operativo.
- c. Identificar la cadena de valor.
- d. Definir el ciclo de vida de los datos.

2. Ejecución:

- a. Estructurar programas y proyectos¹⁵.
- b. Desplegar estrategia de cultura organizacional.
- c. Mapa de ruta.

3. Seguimiento y control:

- a. Indicadores de seguimiento y control.

Cada uno de los elementos descritos anteriormente, se detallarán en capítulos independientes, partiendo desde una estructura general, hacia detalles más específicos; y para efectos del planteamiento de los programas y proyectos se

¹⁴ Concepto definido en el apartado 2.4, tener en cuenta que el gobierno de datos, se enmarca dentro de la estrategia de gestión de datos; el primero no puede existir sin el segundo.

¹⁵ Hace referencia a las disciplinas, componentes y/o áreas de conocimiento.

realizarán ejemplos prácticos que permitan visualizar la aplicación de los conceptos y formatos propuestos, dentro del contexto de las CCF.

4.1 Factores de éxito

Como en toda iniciativa nueva dentro de las organizaciones, es necesario y resulta fundamental asegurar una lista de ingredientes que serán claves en el éxito de los propósitos planteados.

Dichos elementos limitan el quehacer actual y el futuro al que se pretende llegar, se fija el alcance para separar los programas y el cronograma para desarrollarlos. Constituyen el primer paso para ejecutar un correcto proceso de planeación del GD.

Los enunciados de los factores de éxito obedecen a principios generales, los cuales son un punto de partida para adaptarlos a la gestión propia de las necesidades y brechas identificadas por cada organización en el modelo de madurez.

Adicional al respaldo directivo¹⁶, se proponen los siguientes:

Misión

El gobierno de datos es:

- Articulador de la gestión y administración de datos e información de la CCF, el cual se encarga de generar valor en todo el ciclo de vida.
- Marco de referencia válido para la toma de decisiones e identificación de responsables.

Visión

El gobierno de datos será:

- Habilitador de todas las iniciativas relacionadas con los datos e información de la CCF.
- Orquestador para una adecuada exploración y explotación de datos.
- Soporte para la generación de ingresos, identificación de nuevos negocios y aumento de la productividad.

¹⁶ Para efectos de este ejercicio, se da por establecido que el respaldo a nivel directivo se encuentra resuelto, ya que una iniciativa de esta embergadura no puede ser definida ni implementada, sin el compromiso del cuerpo directivo. El modelo de GD implica cambios en los procesos, personas e infraestructura que no se pueden llevar a cabo, sin un respaldo de la dirección.

Objetivos

El gobierno de datos realizará:

General:

Coordinar personas, procesos y tecnología mediante programas, proyectos y estrategias para soportar los objetivos estratégicos de la Corporación.

Específicos:

- Definir el modelo operativo de gobierno de datos de la organización.
- Establecer esquema de calidad y auditoría de datos.
- Identificar responsables en el ciclo de vida.
- Idear estrategia de comunicación previa y durante la implementación.
- Analizar la cadena de valor y ciclo de vida de los datos.
- Implementar las disciplinas y áreas de conocimiento priorizadas (datos maestros, metadatos, glosario de negocio, arquitectura, seguridad, entre otros).

Alcance

El programa de gobierno de datos alcanzará de forma gradual a todas las áreas de la CCF, pero en tres niveles diferentes:

- Áreas productoras de datos que generan ingresos operacionales e ingresos por ganancias (Aportes, Supermercados, Medicamentos, Salud, Vivienda, Crédito, Recreación, Turismo, Educación, Capacitación, Convenios).
- Áreas productoras de datos que no generan ingresos (Mantenimiento, Infraestructura, Compras, Logística, Servicios Administrativos).
- Áreas productoras de datos que operan como apoyo (Auditoría Interna, Revisoría Fiscal, Oficina Jurídica, Planeación, Innovación).

Impulsores

Las razones de iniciar la iniciativa del GD son:

- Preparar a la organización para la adquisición de nuevas competencias, en una economía basada en datos.
- Custodiar la exploración y explotación de los datos en pro de identificar nuevos negocios e incremento de los ingresos.
- Acompañar programas de mejora tecnológica, velando para incluir los elementos de control necesario.
- Definir roles y responsabilidades en los pasos del ciclo de vida de los datos.
- Potencializar los proyectos derivados como glosario de negocio, calidad, seguridad, datos maestros, metadatos, etc.

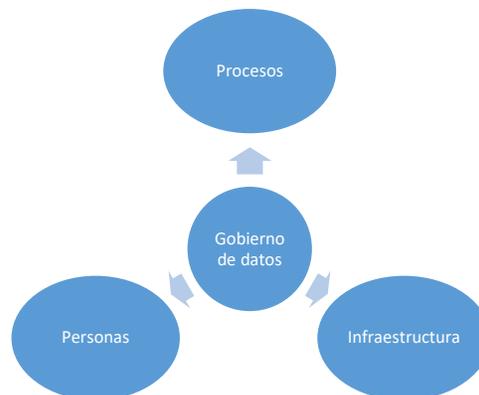
- Crear cultura del dato en la organización para la toma de decisiones.

Capacidades

Las capacidades son un pilar fundamental del modelo de gobierno de datos (ver ilustración 12), ya que son los habilitadores que permitirán el adecuado desarrollo del GD, constituyen nuevos desafíos al tener que generar una evolución en la forma tradicional de gestionar los datos.

Dentro de los programas, proyectos o áreas de conocimiento se encuentran las capacidades de forma implícita, ya que los procesos, la infraestructura y las personas son transversales para cada iniciativa; de ahí la importancia de lograr una adecuada articulación para garantizar el éxito.

En el capítulo anterior se trabajó el diagnóstico por medio de la implementación del nivel de madurez y se logran identificar brechas o necesidades que posee la organización, las cuales se traducen en alguna de las tres dimensiones acá tratadas; por tal motivo, éstas se convierten en el insumo para identificar oportunidades que se traducen en planes, programas o proyectos.



*Ilustración 12 Pilares del modelo de gobierno de datos
Fuente: elaboración propia*

- Procesos

Articulación de las etapas que surten los datos para su producción y consumo. Esta capacidad se involucra desde los sistemas transaccionales, hasta el posterior uso; en la medida que los métodos estén alineados con la estrategia del gobierno, éste se podrá adelantar de forma adecuada.

Es importante tener un área o responsable independiente para dirigir esta capacidad.

Para este ejercicio se propone adelantar dos actividades para asegurar esta capacidad. En primer lugar, un levantamiento inicial de los procesos (tabla 1), para hacer un análisis de brechas con el estado deseado, el objetivo es establecer qué se encuentra identificado actualmente para realizar el empalme con las nuevas exigencias de datos.

Y, en segundo lugar, la aplicación de formato estándar con enfoque hacia los datos (tabla 2), cuyo propósito es detallar y estructurar una metodología que permita recolectar los procesos nuevos; para efectos prácticos se realiza un ejemplo que permite entender el formato, ya que es necesario definir la estrategia, el proceso, el objetivo, el público impactado y por supuesto, las actividades a desarrollar dentro de este bosquejo.

Tabla 1 Identificación inicial de procesos

Consecutivo	Inventario de procesos actuales	Estado futuro
PGD-01	Generación de reportes legales	Garantizar que los procesos actuales estén asegurados y alineados con la estrategia de gobierno de datos.
PGD-02	Gestión de accesos (roles y perfiles)	
PGD-03	Gestión de proyectos	
PGD-04	Gestión de arquitectura de TI	

Tabla 2 Formato levantamiento de procesos

CCF:	Caja de las Oportunidades		
Estrategia:	Gobierno de datos	Versión: 1.1	Consecutivo: PGD-05
Macro-proceso:	Planeación		
Proceso:	Aprobación de iniciativas (programas y proyectos)		
Área:	Comité de gobierno de datos	Gestión de la información	
Objetivo:	Establecer flujo de actividades para aprobar las iniciativas de gestión de datos		
Definición:	<ol style="list-style-type: none"> 1. Recibir documento por medio de un representante del comité o cualquier grupo de interés. 2. Analizar propuesta de acuerdo a calificaciones de impacto. 3. Determinar roles y responsables. 4. Definir cronograma. 5. Realizar seguimiento. 6. Firmar acta de cierre. 		
Responsable:	Secretario del comité de gobierno de datos		
Áreas interesadas	Auditoría Interna, Oficina Jurídica, Tecnología, Negocio		
Recursos	Personas o tecnología involucrados dentro del levantamiento de procesos.		
Entregables	Versiones entregadas y aprobadas por el comité.		
Mediciones	Indicadores de seguimiento que permitan monitorear la iniciativa.		

Los procesos que se podrían identificar serían: aprobación de iniciativas, gestión y control, definición de roles y responsabilidades, estructuración de políticas, estrategia de calidad, estrategia de seguridad, estrategia de metadatos, estrategia de arquitectura, de glosario de negocio, extracción, explotación y exploración de datos, cadena de valor, ciclo de vida, entre muchos otros.

- Infraestructura tecnológica

Es responsabilidad de cada uno de los frentes de trabajo, garantizar que se cuenten con todos los recursos tecnológicos y los motores que se encarguen de asegurar la calidad, la captura y el procesamiento de los datos.

Ante la diversidad de servicios de las CCF, es necesario disponer una bodega¹⁷ única de datos, la cual permitirá tener una mirada unificada del cliente desde la afiliación hasta el uso de los servicios, es un ejemplo claro del análisis a realizar en este componente.

Se propone realizar un listado con el inventario de elementos tecnológicos existentes, para analizar la pertinencia de mantenerlos, modificarlos o cambiarlos; y garantizar las herramientas para el gobierno y la gestión de datos. En la tabla 3 se presenta un ejemplo para ilustrar el enunciado.

Tabla 3 Identificación inicial de artefactos tecnológicos

Inventario de artefactos tecnológicos	Estado actual
Oficina de tecnología	Existe y gestiona toda la operación de TI.
Todos los negocios cuentan con sistemas transacciones	Se presentan ventas manuales, que se deben automatizar.
Reportes automáticos	Se realizan conversiones manuales que dependen de un cargo y no del sistema.
Datos integrados	No existe integración de datos.
Repositorio / bodega de datos	Solo algunos datos se encuentran en el repositorio.
Arquitectura de datos	Existe parcialmente procesos de arquitectura definidos.
Oficina de seguridad de la información	No existe.
Repositorio / bodega de datos maestros y referenciales y metadatos	No se han definido.

De acuerdo al nivel de madurez se determina si la organización está preparada o no con todos los medios necesarios en el componente de TI, para lograr ejecutar de

¹⁷ La unificación de los datos permite acceder y democratizar su uso, tanto para propósitos legales, comerciales y estratégicos.

manera acertada las actividades planeadas; las TI corresponden al soporte técnico que garantizan la aplicación de la iniciativa.

- Personas

Es tal vez el componente más sustancial a considerar en la planeación e implementación; ya que ahí reside el conocimiento tácito y explícito de la organización y se requiere identificar las habilidades necesarias para su desarrollo, así como el aseguramiento de la comunicación y despliegue del GD.

En la nueva economía, las corporaciones dependerán más que nunca de las personas y su capacidad de innovación y así lograr nuevas ideas que transformen y generen valor (Canals, 2004); en este sentido la estrategia de GD se cimienta en las personas al interior de la organización y su capacidad de asimilar la iniciativa como parte de su labor diaria y no como una tarea adicional.

Para analizar este pilar al interior de la organización, se identifican las áreas involucradas con el quehacer de los datos, para asegurar los cargos claves, aplicando la metodología de la tabla 4.

Tabla 4 Identificación inicial de áreas y personas

Entregable	Área	Responsable	Experiencia / Formación	Plan de cierre de brechas
Reporte de afiliados a la CCF	Gerencia Operación de Subsidios	Coordinador Procesamiento de datos	10 años en el cargo, maestría en Ciencia de Datos	No aplica
Reporte afiliados por segmento	Gerencia de Inteligencia	Auxiliar de bases de datos	6 meses en el cargo, tecnólogo en administración	Se incluye dentro del programa de cierre de brechas.
Coberturas de Recreación y Educación	Gerencia de Servicios Sociales	Auxiliar administrativo y financiero	2 años en el cargo, profesional en estadística	Reforzar conceptos y herramientas de análisis de datos.

En la tabla mencionada, se relaciona por cada uno de los entregables actuales -el área y el responsable-, para identificar la experiencia y la formación de las personas y poder aplicar una actividad de plan de cierre de brechas en el caso de aplicar. Al final se obtendrá un inventario que permite analizar las capacidades actuales de las personas y apuntar a su evolución.

El objetivo es identificar las audiencias en cada uno de los pasos del ciclo de vida de los datos, apuntando a nombrar usuarios clave dentro del proceso y de esta manera lograr la movilización requerida; esto porque el GD no lo hace un cargo en específico, por el contrario, se convierte en una labor de todos independientemente del rol dentro de la organización.

Políticas

Constituyen las reglas fundamentales para el proceso de gestión y administración de los datos y las actividades complementarias:

- Debe existir un programa continuo de GD, liderado por un área corporativa adscrita a la Dirección Administrativa.
- Establecimiento de bodegas y repositorios únicos de datos, para garantizar fuentes únicas de verdad.
- Se creará y mantendrá un comité de GD, alineado con todas las áreas responsables.
- Actualización permanente de los roles y responsabilidades.
- Adecuación de flujo y toma de decisiones para los temas relacionados con los datos.
- Creación de un plan de comunicación específico al inicio y durante el despliegue de la iniciativa de gobierno.
- Desarrollo y actualización de las capacidades (procesos, infraestructura y personas).

Ahora bien, una vez declaradas las políticas, el reto más importante es lograr poner en operación los enunciados, de tal manera que se traduzcan en acciones concretas y aplicables a la empresa, se propone el siguiente modelo:

Tabla 5 Fases de operación de políticas

1. Declaración	2. Identificación	3. Asignación	4. Definición	5. Formalización y seguimiento
Política definida, alcanzable y acorde a las necesidades de la empresa.	Roles y responsabilidades asociadas a la política definida.	Identificación de capacidades impactadas (procesos, personas y tecnología)	Entregable a generar de acuerdo a la política.	Entrega del documento soporte de la política. Actividades de seguimiento de las actividades planteadas en el cronograma.

Tal como se puede apreciar en la tabla 5, la puesta en marcha de las políticas debe empezar con la declaración clara y alcanzable, para llegar a una fase de formalización y seguimiento.

De acuerdo con Soares (2014), se requieren una serie de ingredientes necesarios para garantizar el éxito del DG:

- Empoderamiento del negocio con fuerte apoyo de TI: entender que el GD es del y para el negocio, el área de tecnología sólo es el soporte que apalanca la infraestructura técnica.

- Focalización en datos y artefactos¹⁸ críticos: ante la alta cantidad de temas que hacen parte de la agenda del GD, es necesario realizar priorización de las iniciativas, centrándose en aquellas que derivan en un riesgo inminente.
- Alineación de las métricas con las políticas: incluye los mecanismos que permiten desarrollar las métricas, de acuerdo a las definiciones de negocio.
- Victorias tempranas: haciendo referencia a desplegar este tipo de programas paso a paso, ejecutando de forma integral todos los medios necesarios, pero períodos de tiempo más cortos.

Así las cosas, los factores que garantizan de éxito del GD, se convierten en aseguradores para definir hacia donde se enfocan los esfuerzos y las capacidades para su establecimiento.

4.2 Modelo de operación

El modelo de operación es fundamental para el esquema de GD en una CCF, es la muestra cómo operar y gestionar los datos; dicho modelo se visualiza desde los datos, el GD y la estructura organizacional como tal.

4.2.1 De la gestión operativa de datos

Por la diversidad de servicios de las CCF, los sistemas transaccionales también lo son. Tal como se muestra en la ilustración 13, en el centro se encuentra la entidad de afiliado o no afiliado con todos los sistemas transaccionales que se requieren para la venta de productos o prestación de servicios; esto se debe porque la forma de operar cada sector en el que participan las CCF es diferente, no es lo mismo hablar de una reserva de hotel, a una cita médica o a una transacción en un supermercado o droguería.

¹⁸ Se define artefacto como glosario de negocio, reglas de negocio, políticas, estándares, procesos, herramientas e incidentes de calidad, entre otros



Ilustración 13 Operación sistemas transaccionales CCF
Fuente: ilustración propia

En este sentido, se propone crear un único repositorio de datos en el cual se vinculen los datos generados desde el proceso de afiliación hasta la prestación de los servicios. De esta manera se garantiza la definición del Golden Record, ya que se aprueba una única fuente de verdad, el objetivo es minimizar el riesgo asociados a los datos (valores distintos para la misma variable, calidad no adecuada, elevados costos, entre otros).

El detalle de la propuesta de la gestión técnica operativa de los datos se muestra de la siguiente manera:

Gobierno de datos					
Fuentes de datos	Ingesta y mapeo de datos	Catalogación	Calidad	Explotación y explotación	Resultados
	Cargar datos	Modelo	Limpieza	Modelos analíticos	
	Descripción de datos	Relaciones	Reglas de calidad	Reportería	
		Glosario de negocio	Reglas de negocio	Marketing	
Responsable	TI	TI – Negocio - Gestión de la información		Negocio	

Ilustración 14 Gestión operativa de datos
Fuente: ilustración propia

En la base se visualiza los responsables por cada pilar definido, bien sea el área de TI o el negocio de la CCF, con la dirección constante del GD.

La fuente de datos, hace referencia a todos los sistemas transaccionales que posee la corporación, por medio de los cuales realiza la captura, almacenamiento, transformación y consumo de los datos. Constituyen el insumo para la operación y producción del gobierno.

Los 4 pilares propuestos para generar los resultados de los objetivos planteados, se definen así:

- **Ingesta y mapeo de datos:** se incluyen aquellas actividades de cargar y describir los datos que se usarán para un propósito determinado (seguimiento interno, reporte legal, estrategias comerciales). Proceso técnico en el que se realiza la migración de los datos desde los sistemas transaccionales, con el propósito de unificar y centralizar. En este primer paso el objetivo es realizar una aproximación inicial al conocimiento de los datos.
- **Catalogación de datos:** comprende el modelado, el glosario de negocio y estudio de las relaciones de los datos, los cuales se abordan desde el punto de vista tecnológico y de negocio. Estos componentes se desarrollarán en mayor medida en las áreas de conocimiento.

Vale la pena definir que el modelado de datos de acuerdo a la guía de términos de la DAMA (2017) es la representación abstracta de como algo trabaja o está construido, y el diccionario de datos y glosario de negocio, son aquellos conceptos que definen los datos desde el punto de vista técnico – transaccional y de negocio.

- **Calidad de los datos:** está enmarcada desde la limpieza y las reglas establecidas en los sistemas transaccionales y en el repositorio; apuntando a los registros duplicados, inconsistentes, vacíos, etc. La calidad se debe abordar como un caso de uso de tal manera que se aplique sobre un dominio o con un conjunto de dominios para ser asertivo en el desarrollo del esquema.

Para Soares (2010) la definición de la calidad es aquella donde se tienen datos que sean adecuados para su uso. Al igual que en el pilar de catalogación, este componente se detallará dentro de las áreas de conocimiento.

- **Exploración y explotación:** abarca aquellas actividades tendientes a generar modelos analíticos que permitan identificar patrones para mejorar los procesos internos, generación de valor al cliente, nuevas oportunidades de negocio, etc.; cumpliendo así mismo con las disposiciones legales vigentes que existen sobre la operación de las CCF. Incluye los modelos analíticos, de reporte y de marketing

Por lo tanto, los resultados del modelo operativo de los datos son analizados desde los objetivos planteados, es decir, desde la generación de valor para el cliente o para el mejoramiento de los procesos.

4.2.2 Del comité de gobierno de datos

El modelo de operación del comité de GD a alto nivel se define como un proceso único y articulador entre el área de TI, áreas de apoyo y el negocio; como organismo rector y regulador realiza los esfuerzos para mantener bajo su sombrilla las capacidades de todos los actores con los procesos estratégicos, tácticos y operativos de la CCF.

El comité de GD cobra sentido cuando cumple las funciones de regir y direccionar todos los temas asociados a la gestión de datos de las CCF.

A nivel de integración desde el punto organizativo, los responsables y representantes se presentan en la ilustración 15; se enuncian los claves por áreas, pero pueden existir otros de acuerdo a los dominios que se están trabajando.



*Ilustración 15 Modelo operativo del comité de gobierno de datos
Fuente: ilustración propia*

Responsable del gobierno de datos: garantizar el cumplimiento de los planes, programas y proyectos inherentes al GD, así como la definición de las políticas y los procesos.

Representante de tecnología: proporcionar la infraestructura y arquitectura que garantiza la implementación del GD.

Representante de auditoría interna: velar porque los procesos se ejecuten de acuerdo a las políticas definidas, con el objetivo de minimizar los riesgos asociados.

Representante de la oficina jurídica: asegurar que el marco normativo se encuentre presente en todas las iniciativas dirigidas por el comité de GD.

Representante de Negocio:

- Líder estratégico: hace referencia al rol directivo del negocio donde se concentra la toma de decisiones propia de la operación. Es el aliado para garantizar la ejecución y cumplimiento del nivel táctico y operativo.
- Líder táctico: se convierten en los inductores para la ejecución de planes de acción, así como de calidad de datos, y también como propietarios de los datos identifican nuevos campos de trabajo.
- Líder operativo: ya que son responsables operativos de los datos, ejecutan las actividades determinadas dentro de la estrategia de gestión de datos.

Los macro-procesos y actividades que el comité del GD adelanta, se especifican en el capítulo 4.3 Cadena de valor.

4.2.3 De la estructura organizacional

Con la definición del modelo de operación de los datos y del comité, el siguiente estadio es acordar la estructura organizacional ideal, la cual debe ser ajustada a la gestión y control de los datos e información.

Como ya se ha mencionado, no hay una fórmula única que se aplica como una receta única, por el contrario, se utilizan puntos de referencia que se adaptan a la organización, se mejoran cuando los procesos evolucionan y se adaptan a los cambios de las sociedades y los trabajadores del conocimiento.

Lo importante es incluir roles y responsables visibles dentro de la organización, los cuales realizan las actividades de CDO (Chief Data Officer); de acuerdo con Soares (2014), el nacimiento de un CDO se debe a:

- Fuerte ambiente regulatorio en torno a los datos.
- Obligación de valoración o monetización de los datos.
- Crecimiento exponencial de los datos, así como el volumen y variedad (Big Data).
- Políticas de gestión interna.

En cuanto al nivel de centralización se propone un CDO a nivel Corporativo, el cual trabaja de forma transversal para toda la CCF y un Manager¹⁹ para cada una de las áreas de negocio.

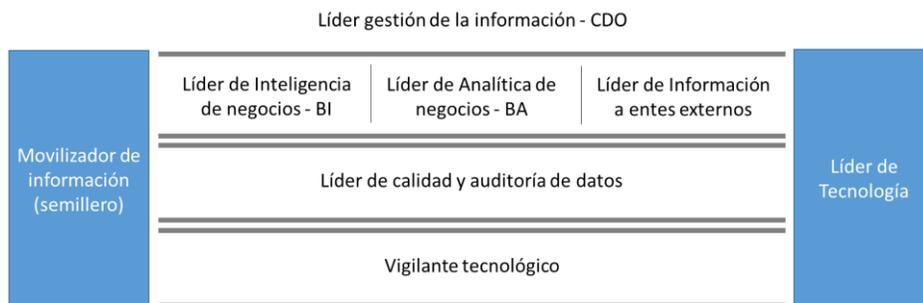
Lo anterior se propone porque al ser empresas de tipo multi-negocio, se requiere un proceso integral y holístico que garantice la unicidad de los datos generados en las áreas, sin desconocer que los expertos de negocio se encuentran en la parte

¹⁹ Cumple con un rol similar al CDO y está sujeto a las directrices corporativas, se convierte en el aliado desde la operación de las áreas generadoras de ingresos.

operativa, los cuales tienen necesidades de GD muy específicas; sin desconocer que el nivel de Manager depende funcionalmente del CDO Corporativo.

A. Estructura organizacional corporativa

El modelo organizacional desde el punto de vista corporativo se muestra en la siguiente ilustración:



*Ilustración 16 Estructura organizacional nivel corporativo
Fuente: elaboración propia*

Líder gestión de la información - CDO: dentro de la agenda de este rol articulador se encuentran las siguientes actividades:

- Construir la estrategia de gestión de datos de la Corporación.
- Liderar estrategia de gobierno de datos, nombrando al líder y a los responsables.
- Establecer la línea de autoridad para la toma de decisiones en los temas relacionados con los datos.
- Identificar los roles claves dentro de la organización para articular trabajo unificado.
- Articular políticas, estándares y procesos para garantizar el correcto despliegue de las propuestas de datos.
- Acordar las responsabilidades de los dueños y custodios de los datos.
- Modular estratégicamente las aplicaciones, repositorios, bodegas y conjunto de datos en la Corporación.
- Alinear a los propietarios con el uso de los datos, para potenciar su desarrollo.
- Obtener junto con los propietarios, los niveles de confianza de los datos bajo su responsabilidad, garantizando la seguridad y privacidad de los mismos.

Líder de inteligencia de negocio – BI: las funciones para este frente de trabajo se centran en:

- Aplicar metodologías y herramientas para el análisis de datos.
- Planificar y desarrollar los indicadores de gestión definidos de acuerdo a la necesidad del negocio y del análisis propuesto.
- Generar soluciones estadísticas para el análisis de información, garantizando la adecuada exploración y explotación de los datos.

- Realizar las actividades propias de creación de reportes y tableros de control, que permitan la visualización de información, con las herramientas habilitadas por la Corporación.

Líder de analítica de negocios – BA: el alcance para este rol se centra en:

- Generar modelos analíticos de tipo descriptivo, predictivo y prescriptivo con información relevante y estratégica.
- Aplicar herramientas de minería de datos para análisis cualitativos y cuantitativos.
- Apalancar las estrategias comerciales, de definición de clientes, productos, mercados objetivos.
- Estructurar modelos del sector de las CCF, con estudios que provean información relevante para el sistema de Compensación Familiar.

Líder de información a entes externos: se propone un área especializada que se encargue de dar respuesta a los requerimientos de tipo legal, las cuales se resumen así:

- Generar los reportes a entes externos de acuerdo las directrices reglamentadas por dichas entidades.
- Garantizar la entrega con los contenidos mínimos de oportunidad y calidad.
- Orientar y definir los criterios los reportes a entregar, aplicando en conjunto con el área de conocimiento de calidad de datos, las reglas necesarias para reducir a cero el riesgo asociado.

Líder de calidad y auditoría de datos: con el propósito de garantizar la aplicación adecuada de la estrategia de calidad de datos, resulta indispensable visualizar un cargo que de dedicación exclusiva a esta labor:

- Idear estrategia de calidad, así como su correcto despliegue.
- Identificar roles y responsabilidades asociados a la calidad, aplicando las dimensiones que se consideren necesarias.
- Estructurar indicadores de seguimiento para la correcta adopción del GD.
- Definir las reglas iniciales, así como el acompañamiento para la identificación de nuevas reglas, las cuales se aplican en los niveles definidos.

Vigilante tecnológico: ante el crecimiento y desarrollo de TI constante, se requiere un responsable de monitorear el entorno para la potencialización de los datos de la Corporación:

- Analizar el mercado, tendencias y nuevas formas de administrar, controlar y gestionar los datos.

- Proponer nuevos modelos de análisis y explotación de datos, de acuerdo a la información sectorial.
- Desarrollar estrategias para establecer comparativos con empresas homólogas.

Movilizador de datos (semillero): el movilizador es un rol que, ejerciendo cualquier cargo dentro de la compañía, se convierte en un aliado para el GD, tiene interés particular en el tema y se transforma en el aliado perfecto para la operación e implementación del programa.

- Participar activamente en las actividades de gestión de cambio de la estrategia de GD
- Realizar las capacitaciones y demás mecanismos que permiten adherir la gestión de los datos, como una capacidad más dentro de la CCF.
- Promover y difundir el plan estratégico del GD.

El semillero más allá de un cargo, se convierte en un esquema en el que se formarán y desarrollarán las capacidades de los profesionales en torno a la gestión de los datos y de la información, son los futuros trabajadores del conocimiento.

Este programa se convierte en una comunidad de interés, en la cual se analizan y debaten los temas inherentes a los datos, y de esta manera encontrar puntos de evolución, de control y gestión; lo cual permitirá que el GD se visibilice en todos los niveles de la organización.

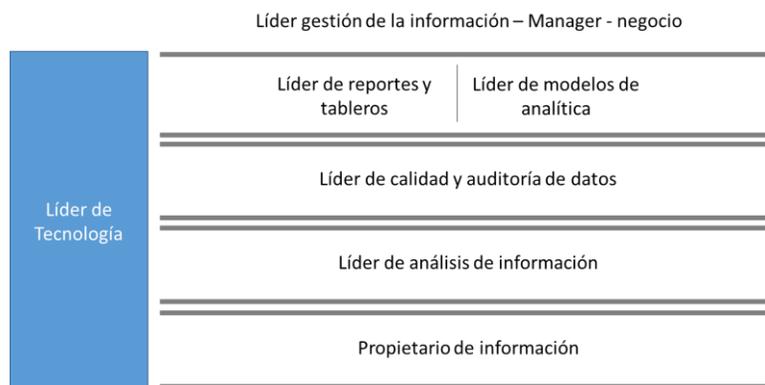
Tecnología:

- Garantizar los habilitadores digitales y arquitectónicos para implementar los planes, programas y proyectos.
- Proponer nuevas metodologías de control y gestión de los datos, de acuerdo a la estructura del repositorio de información definido.
- Proveer las herramientas necesarias para garantizar la custodia de los datos, así como la seguridad y privacidad.

Vale la pena recordar que los roles acá presentados, son un punto de partida para la adaptación a las condiciones financieras y organizacionales de cada CCF, el objetivo es garantizar que las funciones planteadas sean absorbidas por un responsable de su ejecución. Igualmente, la cantidad de personas necesarias para cada frente depende del análisis de tiempos y movimientos que realice la capacidad de procesos.

B. Estructura organizacional de negocio

Debido a que los datos se producen desde los sistemas transaccionales en las unidades de negocio, es necesario incluir una estructura de datos a nivel organizacional para cada una de las áreas productoras de datos, para garantizar de esta manera una explotación, exploración y custodia de los datos (ilustración 17).



*Ilustración 17 Estructura organizacional a nivel operativo de negocio
Fuente: elaboración propia*

Las funciones son homólogas a las presentadas en la propuesta transversal, con la diferencia que en este caso el campo de acción corresponde al negocio. Únicamente se presentan tres cargos nuevos:

Líder de reportes y tableros:

- Generar los reportes necesarios para el monitoreo del negocio a cargo.
- Estructurar indicadores de seguimiento de negocio de variables operativas para gestión diaria.
- Identificar necesidades de visualización de tableros que mejoren la gestión de la información.
- Aplicar y operar las herramientas y directrices impartidas desde el área corporativa en torno a la gestión de datos.

Líder de análisis de información:

- Aplicar las técnicas estadísticas para el análisis de información del área a cargo.
- Construir información estadística que genere valor dentro de la empresa.
- Identificar patrones de comportamiento y de tendencias, para la generación de nuevos ingresos.
- Contribuir con sus análisis al desarrollo y madurez de los modelos analíticos establecidos.
- Desarrollar nuevos análisis de acuerdo a los pilares de la exploración y explotación de datos.

Propietario de información:

- Ejercer supervisión para garantizar el cumplimiento de procesos, procedimientos y políticas definidos.
- Garantizar que la operación diaria del gobierno de datos se realice, traduciendo los requerimientos de negocio al área de tecnología.
- Realizar tareas de retroalimentación en los diferentes niveles de la organización, en temas referentes al gobierno de datos.
- Identificar insumos para plantear iniciativas en torno al desarrollo de los datos de operación de negocio.
- Controlar los datos producidos a su cargo, para responder por la producción y el consumo en el ciclo de vida.
- Participar activamente en todas las iniciativas asociadas a la gestión de los datos, para garantizar su desarrollo.

Dentro de la literatura existente, se habla de establecer una oficina de gobierno de datos de forma independiente, lo cual depende de la capacidad administrativa y financiera de la CCF; en esta propuesta se incluye esta función en el espectro de la agenda del CDO, cuyo objetivo es garantizar la integración de los elementos del gobierno con los responsables de cada área y expertos de negocio.

4.3 Cadena de valor

Tomando como referencia el concepto de cadena de valor (CV) de Michael Porter (1985), en este apartado se aborda una aproximación para generar los macroprocesos que permitan identificar aquellas actividades estratégicas y generar el valor real de los datos. Las actividades tradicionales que las empresas identifican como claves, ya no son las mismas en las sociedades del conocimiento; por lo tanto, el proceso de toma de decisiones hoy en día se realiza basado en los datos (Muñoz, 2014).

4.3.1 De la organización basada en datos

Por el desarrollo tecno-científico de las economías, las empresas implementan métodos de cambio y adaptación a esos nuevos retos; por este motivo es transcendental modificar la forma de operación, pasando de un concepto tradicional de tenencia física, a un concepto digital y por lo tanto de datos; es decir, evolucionar hacia una corporación manejada con datos.

Las CCF, no son ajenas a esta realidad y se deben emitir actividades para transformarse y adaptarse a este nuevo entorno. Resulta imperioso adecuar las cadenas de valor a un ecosistema donde la competencia la gana quien mejor administre y gestione los datos.

En la ilustración 18, se propone un modelo de cadena de valor para una CCF, gestionada y gobernada por datos.



Ilustración 18 Modelo de una organización manejada por datos
Fuente: elaboración propia a partir de Muñoz y Quiroz (2014)

Actividades primarias: aquellas fundamentales para la generación de valor de las CCF.

- Activos: dentro de los conceptos de la DAMA (2011) del diccionario de la gestión de datos se hacen las siguientes definiciones:

Dato: hechos representados como texto, números, gráficos, imágenes, sonido o video.

Información: comprensión relativa a cualquier objeto, como hechos, eventos, cosas, procesos o ideas. Conceptos que dentro de un determinado contexto y período de tiempo tienen significado particular.

Conocimiento: comprensión del significado de la información. Información en perspectiva, basada en el reconocimiento de patrones y la experiencia.

- Ciencia de los datos: de acuerdo a las definiciones de Oracle²⁰, se trata de obtener el valor de los datos a través de la integración de disciplinas (estadística, matemáticas, programación, etc.) en un solo proceso, incluyendo el conocimiento de negocio.

Metodologías: para estructurar estrategia de gestión de datos, de gobierno de datos.

²⁰ Recuperado de: <https://n9.cl/kc5a>

Métodos: para realizar análisis descriptivo (indagar el pasado), predictivo (pronosticar y predecir), prescriptivo (detectar mejores alternativas) y cognitivo (generar conocimiento, aprender).

Herramientas: para crear y administrar repositorio de datos, bodega de datos, visualizadores de reportes y tableros, modelos.

Actividades de apoyo: necesarias y de importancia absoluta para garantizar que la organización sea administrada bajo el concepto de datos, se convierten en habilitadores para que las actividades primarias puedan ejercer su labor de forma adecuada.

- **Capacidades:** este apartado se describió en el capítulo 4.1, destacando que las personas, los procesos y la tecnología están al servicio del GD, y funcionan de forma alineada y articulada para garantizar el éxito.
- **Cultura:** este eslabón juega un papel importante dentro del despliegue de toda iniciativa empresarial, y la estrategia de GD no es la excepción. Aplicando las decisiones que permitan la adopción del programa, se podrá llegar a todos los públicos definidos y de esta manera garantizar que los datos, sean parte del ADN empresarial a todo nivel. En el capítulo 4.6 se realizará un análisis más detallado de cómo es la definición de un proceso cultural para GD.

Así las cosas, los datos se han convertido en el eje central de los modelos de negocio de las empresas exitosas, y no pueden ser un camino desconocido para las CCF. Por tal motivo, resulta necesario definir una cadena de valor empresarial que se ajuste a un nuevo entorno y prepare a la Corporación para ser más exitosa en un mundo competido teniendo a los datos como factor diferenciador.

4.3.2 Del comité de gobierno de datos

Los macro-procesos del comité de GD como entidad reguladora y tomadora de decisiones dentro del despliegue y la continuidad, se visualiza en la ilustración 19:



*Ilustración 19 Macro-procesos del comité de gobierno de datos
Fuente: elaboración propia.*

El objetivo es identificar aquellas actividades claves que generen valor dentro del rol que desempeña el comité, en todo el proceso de la gestión de datos de la CCF.

Planeación: macro-proceso encargado de gestionar el gobierno de datos estratégicamente, definiendo el plan de trabajo y articulación de todos los participantes.

- Definición: plan de trabajo de implementación del GD, seguimiento a las actividades planeadas, determinación de responsables.
- Investigación: mejores prácticas de mercado que permitan garantizar la estabilidad, desarrollo y continuidad del GD.

Políticas, procesos y procedimientos: incluye la definición, aprobación y gestión de la documentación y directrices impartidas desde el gobierno.

- Calidad de datos: generar las definiciones con los requerimientos necesarios para el análisis de causa raíz, así como la aprobación de las reglas definidas.
- Desarrollo de los datos: propuestas en torno a identificar programas y proyectos que permitan la evolución del modelo y avanzar hacia el objetivo de gestionar la CCF basada en datos. De acuerdo al diccionario de términos de la DAMA (2011).
- Priorización de iniciativas de datos: en atención a la diversidad de necesidades, programas y proyectos que se presentan para la gestión de datos, el primer reto es planear que hacer primero.

Como herramienta de apoyo se propone una matriz que permita realizar la medición de la priorización (tabla 6), la cual se describe así: de acuerdo a cada iniciativa²¹ presentada se mide el impacto estratégico, económico y legal (de 1 a 5), de esta manera se obtiene la puntuación total (de 0 a 15), para llegar a la columna de prioridad (alta, media, baja).

Tabla 6 Matriz de priorización de iniciativas de datos

Iniciativa	Responsable	Gerencia / Área	Alcance	Duración	Impactos ²²			Puntuación total	Prioridad ²³
					Estratégico	Económico	Legal		
Unificar estructura única de nomenclatura de tipo de	Sully Thomson	Gerencia Operación de Subsidios	Afiliados y sistemas transaccionales	1 año	4	3	2	9	Media

²¹ Debe tener un responsable, el área solicitante y alcance.

²² Valoración de 1 a 5, donde 1 es el nivel más bajo y 5 es el nivel más alto.

²³ Resultado de la medición de la valoración, 0-5 alta, 6-10 media, 11-15 alta.

documentos en todos los sistemas transaccionales.									
Definir fuente de datos para los afiliados a la CCF.	Homer Tristan	Gerencia Operación de Subsidio	Afiliaciones	1 mes	5	5	3	13	Alta

En el ejemplo anterior, de acuerdo a la columna de prioridad la segunda iniciativa debe empezar con más prontitud, de acuerdo a la calificación obtenida.

Toma de decisiones: dar el marco metodológico respecto al tratamiento para los procesos de calidad y limpieza de datos, así como el uso de las herramientas digitales necesarias para el funcionamiento del gobierno. Entre otras, están:

- Aprobar reglas de calidad para cada uno de los procesos.
- Decidir respecto a novedades de calidad identificadas en el proceso.

Innovación: se convierte tal vez en el eje de mayor relevancia, el uso de los datos de forma estratégica es lo que realmente agrega valor, para alcanzar la transformación basada en los datos. Resulta significativo estar al tanto de estudios y tendencias que permitan un mejor entendimiento del quehacer del GD.

4.3.3 De la explotación y exploración de datos

Se definen los macro-procesos en la ilustración 20 en pro de generar valor con los datos generados en las corporaciones.



*Ilustración 20 Cadena de valor - explotación y exploración de datos
Fuente: elaboración propia*

En el proceso de planeación se consideran actividades que permitan visualizar cuál debería ser la forma de operar el aprovechamiento de los datos, con un cronograma y responsables claros.

El almacenamiento de los datos toma relevancia, porque la unificación y estandarización de repositorios únicos, permiten la explotación de los datos de una forma eficiente.

La generación de reportes y modelos es la traducción del uso de los datos, para los propósitos que la Corporación decida.

En el último eslabón se encuentran las actividades para la búsqueda de resultados, de acuerdo a los objetivos de la obtención de nuevos ingresos, nuevos negocios o mejoramiento en la productividad, lo que al final es lo que buscan los trabajadores del conocimiento y así obtener la sabiduría empresarial, que es la que permite mejorar la competitividad.

4.4 Ciclo de vida de los datos

Para lograr una adecuada gestión del modelo de gobierno, es necesario establecer el ciclo de vida que tienen los datos dentro de la organización, este se identifica de acuerdo a los sistemas transaccionales y a las fuentes de datos externos existentes. De acuerdo a las recomendaciones del CMMI se proponen los siguientes componentes (ver ilustración 21):



*Ilustración 21 Modelo del ciclo de vida de los datos
Fuente: elaboración propia a partir de CMI*

Crear, adquirir y capturar: proceso de creación de los datos en los sistemas transaccionales, así como la adquisición en caso de requerirse y la captura en la prestación de los servicios.

- Sistemas fuentes internos: los datos provienen de los sistemas de las CCF.
- Datos externos: provienen de sistemas externos como las redes sociales.

Procesar y analizar: actividad para generar proceso de ingesta de datos a la ubicación definida.

- Bodega única de datos: repositorio y única fuente de verdad para todos los propósitos de gestión de datos.

Gestionar y preservar: gestionar los datos para ponerlos a disposición de los interesados.

- Construir y depurar: etapa técnica donde se construye y alimenta el repositorio de datos.
- Reglas de calidad: aplicación de los puntos de control definidos.

Acceder y publicar: proceso mediante el cual se le da uso y por lo tanto valor a los datos.

- Reportes y tableros: a partir de los datos existentes realizar la elaboración de reportes internos y externos, de acuerdo a la demanda actual y a los nuevos requerimientos de visualización de información.
- Publicaciones: generar entregables de información en objetos visuales que sean de interés para todas las áreas de la organización, por ejemplo, boletines semanales donde se muestre la información relevante de la CCF, avisos en la intranet corporativa que permitan visibilizar el impulso que la Corporación está ejerciendo sobre los datos.

Aprovechar, reutilizar y archivar: desarrollo y uso de los datos para apalancar los ingresos e identificación de nuevas oportunidades de negocios. Una vez generado el uso del dato, el siguiente proceso es la reutilización o el archivo, de acuerdo al propósito del análisis realizado.

- Modelos analíticos: identificación de oportunidades mediante el entendimiento del cliente y la segmentación estratégica.
- Productos de acuerdo a preferencias: focalización y personalización de productos y/o servicios para los clientes afiliados, de acuerdo al conocimiento y desarrollo de los datos.
- Decisiones estratégicas: como resultado de la disponibilidad de información, convertida en objetos visuales, el proceso de gestionar la CCF sustentada en los datos será parte embebida del despliegue del GD.

El ciclo de vida de los datos, pretende resaltar el proceso de generación, uso y disposición final, desde una mirada estratégica y no técnica.

Se recomienda usar las herramientas de mapas de calor para establecer los tiempos de archivo de los datos, si los datos tienen una fuerte demanda van a estar disponibles en todo momento, por el contrario, si la demanda es baja, se almacenan donde no haya consumo de recursos.

Ejemplo 1. En la siguiente tabla se muestra el camino que tomaría la afiliación de un trabajador.

Tabla 7 Ciclo de vida para la entidad de afiliado

Crear, adquirir y capturar	Procesar y analizar	Gestionar y preservar	Acceder y publicar	Aprovechar, reutilizar y archivar
Captura del tipo y número de documento del afiliado nuevo	Ingesta del tipo y número de documento del afiliado al repositorio de datos	Validación de tipo y número de documento único en formato y nomenclatura establecido	Cantidad de afiliados por semana	Asignación de segmento poblacional para conocer gustos, preferencias, tendencias.

Ejemplo 2. En la siguiente tabla se muestra el camino que tomaría el acceso al servicio de crédito de un trabajador afiliado.

Tabla 8 Ciclo de vida para asignación de crédito

Crear, adquirir y capturar	Procesar y analizar	Gestionar y preservar	Acceder y publicar	Aprovechar, reutilizar y archivar
Número de crédito aprobado a un tipo y número de documento de identidad de un trabajador afiliado.	Ingesta del número de crédito aprobado a la bodega de datos.	Asociar valor desembolsado al cliente único.	Valor desembolsado por semana.	Modelo de predicción para identificar patrones de uso del crédito asignado.

4.5 Lineamientos, disciplinas, y/o áreas de conocimiento

Una vez realizado el análisis y estableciendo los parámetros de los factores de éxito, el modelo de operación, cadena de valor y ciclo de vida de los datos; se procede a la identificación de las disciplinas a gobernar, son medios esenciales para garantizar que el gobierno y la gestión de datos sea integral a todas las áreas de conocimiento.

Al desarrollar un programa de gobierno de datos se contemplan aspectos únicos y mínimos que garanticen una correcta adopción, el enfoque es integrador para alcanzar la relevancia y alto impacto en el futuro de la organización.

Como punto de partida y siguiendo los marcos metodológicos referenciados, se proponen los elementos a desarrollar de la ilustración 22. Esta situación se deriva

porque no existe una fórmula única, exclusiva o cierta; por el contrario, cada iniciativa de GD es distintiva de la corporación donde se aplica.



*Ilustración 22 Modelo de gobierno de datos desde las disciplinas o áreas de conocimiento.
Fuente: ilustración propia a partir de las metodologías de la sección 2.3*

En este documento se utilizan los términos de componente, disciplina y área de conocimiento de forma indistinta, el propósito es destacar aquellos ámbitos que son importantes para incluir dentro de la gestión de datos y por lo tanto del GD.

Resaltar también que las disciplinas pueden llegar a convertirse en áreas independientes dentro de la organización, dada su importancia, su alcance y la posible materialización de riesgos, se pueden crear roles que se encarguen exclusivamente de implementar una disciplina como una estrategia, bajo la sombra del GD.

Así las cosas, se tratará de ahora en adelante cada una de las disciplinas, como proyectos y programas a implementar, de acuerdo a las necesidades de la CCF y a la priorización adelantada desde el comité de gobierno de datos.

4.5.1 Lineamientos generales

Procesos administrativos y de negocio que dan los parámetros mínimos a considerar dentro del marco de referencia, se convierten en el faro que indican el curso de acción, la toma de decisiones y la definición de las iniciativas asociadas a los datos.

4.5.1.1 Misión, visión, objetivos

Tal como se mencionó en el capítulo 4.1 Factores de éxito, el primer paso para el despliegue de la estrategia, incluye todas las actividades correspondientes a la definición de misión, visión, objetivos; es decir, la razón de ser para establecer el

gobierno sobre los datos, el marco metodológico y las capacidades para lograr los objetivos.

El aseguramiento de los factores de éxito se convierte en el primer paso antes de iniciar a construir los soportes de un modelo de gobierno, porque no se trata de una moda, sino de requerimientos estratégicos de negocio.

Las dimensiones para abordar los factores de éxito pueden considerar la calidad, seguridad, requerimientos legales, necesidad de análisis de información internos, mejorar competitividad, entre otros.

4.5.1.2 Políticas de gobierno

Son todas aquellas actividades generales que entregan los lineamientos traducidos en acciones, para el despliegue y continuidad del GD.

- Identificar los impulsores del GD y de cada estrategia para la gestión de datos.
- Establecer comité de gobierno de datos, como órgano e instancia superior y único para la toma de decisiones.
- Aprobar y publicar el documento de GD como forma de hacerlo extensivo a todos los públicos²⁴.
- Implementar un modelo de operación de una empresa, manejada por datos; incluyendo la estructura organizacional.
- Establecer un esquema de cultura organizacional, asociada a la gestión de datos.
- Fortalecer el despliegue de nuevas iniciativas de datos para la identificación de nuevos negocios e incremento de ingresos.
- Establecer cronograma y factores de seguimiento (metas e indicadores) para la ejecución.
- Priorizar de acuerdo a matriz de criticidad las iniciativas asociadas a la gestión de datos.
- Aprobar procesos y cambios en estructura que se consideren necesarios para garantizar el modelo.
- Investigar tendencias, buenas prácticas y nuevas herramientas que permitan el desarrollo y evolución del GD.
- Velar para que cada proyecto cuente con un proceso formal de documentación, garantizando que se incluyan aquellas experiencias que permitan gestionar el conocimiento.

4.5.1.3 Roles y responsabilidades

²⁴ En el capítulo de cultura se propone la creación de un decálogo del gobierno de datos, el cual es entregado y publicado y se encuentra al alcance todas las personas de la organización.

La implementación del GD implica la identificación de todos los grupos de interés alrededor de los datos, sean como proveedores, custodios, dueños, propietarios, clientes, etc., para el caso de estudio se identifican a nivel macro los líderes a nivel de gobierno, de tecnología y de negocio, lo interesante de este ejercicio es crear las tareas a realizar por cada área, líder o rol de forma individual o colectiva.

A continuación, se muestran ejemplos de los líderes mencionados, se describe la contribución general y se detallan las responsabilidades para cada caso:

Líder de gobierno de datos²⁵: encargado de gestionar la ejecución, garantizar la continuidad y la evolución del GD:

- Garantizar cumplimiento regulatorio interno y externo, operando las decisiones establecidas, traducidas en políticas.
- Alinear a todas las áreas interesadas e involucradas para encender el programa, así como las labores administrativas del comité de GD.
- Trabajar de manera articulada con los propietarios y custodios de los datos, para identificar nuevas oportunidades de mejora.
- Implementar la estructuración y desarrollo del glosario de negocio; así como las disciplinas definidas y priorizadas de calidad, datos maestros, seguridad, entre otros.
- Definir métricas e indicadores de seguimiento.

Líder de tecnología: proporcionar la infraestructura y arquitectura que garantiza la implementación del programa:

- Soportar la ejecución del GD a nivel técnico.
- Aplicar los protocolos de seguridad y privacidad de los datos y la información.
- Garantizar disponibilidad del uso de los datos, requeridos por los proveedores internos y clientes finales.
- Evaluar los sistemas actuales para el cumplimiento de los propósitos.
- Aplicar los estándares de arquitectura de datos e información.
- Desarrollar las nuevas necesidades de reportes, tableros, y en general los datos que se requieran.
- Garantizar la aplicación de la estrategia de calidad de datos, dado el alcance que les corresponda.
- Desarrolla estrategias para reducir problemas de datos, asociados al uso de los sistemas de información.

²⁵ Para este ejercicio, el liderazgo de la responsabilidad del GD se incluyó dentro del CDO, sin embargo, de acuerdo a la disponibilidad de recursos de la CCF, se puede crear un cargo independiente, con el mismo alcance acá definido.

Líder del negocio: por lo general son los custodios²⁶ o propietarios²⁷, es donde se produce por medio de los sistemas transacciones el nacimiento y captura de los datos:

- Operar las definiciones decididas de acuerdo a los análisis realizados.
- Aplicar la estrategia de calidad, garantizando de forma interdisciplinaria que las reglas estén adecuadas.
- Acompañar y operar las estrategias de datos maestros, seguridad, metadatos, entre otras; de acuerdo a su rol.
- Garantizar la administración de los datos que se encuentran bajo su responsabilidad.
- Explotar y explorar los datos de su negocio, y de esta manera lograr potencializar la información a su cargo.
- Identificar y contribuir con su conocimiento a la construcción de oportunidades, programas y proyectos que permitan el desarrollo de nuevas iniciativas asociadas a los datos.

Líder de auditoría interna y Líder de la oficina jurídica, que como se mencionó en el apartado 4.2.2, se encargan de vigilar y garantizar el cumplimiento de las definiciones establecidas desde la gestión de riesgos y acatamiento legal.

A manera de conclusión, los roles y responsabilidades acá definidos se encargan de implementar, continuar y evolucionar el modelo de gobierno de datos, así como la puesta en marcha de las iniciativas asociadas; el desarrollo de las capacidades y áreas de conocimiento, serán su responsabilidad desde la posición de cada interesado.

4.5.1.4 Glosario de negocio

Para realizar unificación de términos y entendimiento único del modelo de negocio, es necesario realizar la definición de un glosario que permita tener un único repositorio por cada entidad definida.

El glosario responde a las necesidades y al contexto de cada área analizada, ya que se pueden presentar distintas definiciones de acuerdo al tipo de dato relacionado.

²⁶ Custodio: para este ejercicio se define como aquel rol / cargo que tiene el acceso al dato, quién se encarga de gestionar, administrar, tratar, guardar y almacenar el dato. Esta definición se realiza desde el negocio y no desde la función de TI.

²⁷ Propietario: corresponde al dueño del dato como tal, quién lo genera, crea o captura; es el responsable desde el negocio.

Contiene al menos los siguientes parámetros:

- Organizacionales: corresponde al área productora del dato, donde se captura y bajo de que área se encuentra responsabilizado; se requiere la gerencia, sección, propietario, custodio.
- Operativos: hace referencia la ubicación dentro del repositorio único de datos, donde se encuentra alojado el dato desde el punto de vista técnico; es decir, bodega y el objeto donde se encuentra alojado la entidad definida.

En este mismo sentido, contar con el mapeo de los datos desde los sistemas transaccionales es fundamental, ya que es necesario saber los datos que están poblando en la bodega / repositorio.

- Conceptuales: como tal la nomenclatura de la entidad, desde la perspectiva de negocio, por lo tanto, hace mención al marco normativo, uso y definición como tal.

En las siguientes tablas se ejemplifica la entidad²⁸ afiliados y subsidio en especie con las sub-entidades²⁹ trabajador dependiente y kit o paquete escolar respectivamente:

Tabla 9 Ejemplo glosario de negocio

Entidad	Sub-entidad	Parámetro	Sub-parámetro	Asignación
Afiliados	Trabajador dependiente	Organizacionales	Gerencia	Operación de subsidio
			Sección	Administración de afiliados
			Propietario	Administración de afiliados
			Custodio	Coord. de procesamiento de datos
		Operativos	Bodega	BW-Afiliaciones
			Objeto	BW-SAP-AFI-01
			Sistema fuente	Transaccional de administración y gestión de afiliados
		Conceptuales	Marco normativo	Ley 21 de 1982
			Uso	Reportes legales, segmentación de clientes, asignación de subsidios, prestación de servicios.
			Definición	Los trabajadores de carácter permanente al servicio de los empleadores previstos en los artículos 7o. y 72 de la ley 21 de 1982, desde el momento de su vinculación y hasta la terminación de la misma.
Subsidio en especie		Organizacionales	Gerencia	Operación de subsidios

²⁸ De acuerdo a la definición de la DAMA (2107), una entidad es algo de interés, un objeto o evento, colección de datos que el negocio considera importante y digno de captura.

²⁹ Se emplea para este ejercicio el concepto de sub-entidad para detonar un nivel adicional de análisis, únicamente con el propósito de agrupar las entidades en conceptos más amplios.

	Kit / paquete escolar	Operativos	Sección	Subsidio monetario
			Propietario	Subsidio monetario
			Custodio	Coord. de procesamiento de datos
		Conceptuales	Bodega	BW-Afiliaciones
			Objeto	BW-SAP-SUB-01
			Sistema fuente	Transaccional de administración y gestión de subsidios
			Marco normativo	Decreto 784 de 1989
		Uso	Reporte legal, conocimiento del cliente	
		Definición	Las Cajas de Compensación Familiar, conforme al artículo 5 de la Ley 21 de 1982, podrán reconocer subsidio familiar en especie, consistente en alimentos, vestidos, becas de estudio, textos escolares, drogas y demás frutos o géneros diferentes al dinero.	

Entre las entidades y sub-entidades que se pueden identificar para una CCF, están:

Tabla 10 Identificación de entidades y sub-entidades para una CCF

Entidad	Sub-entidad	Entidad	Sub-entidad
Afiliado	Trabajador dependiente	Fondos sociales – Fovis	Gastos de administración
	Trabajador independiente		Gastos de sistema de información
	Trabajador pensionado		Subsidios asignados
	Trabajador del servicio doméstico		Subsidios girados
	Trabajador facultativo		Subsidios vencidos
	Madres comunitarias		Subsidios renunciados
	Empresas 4%,		Subsidios restablecidos
	Empresas ley 1429		Unidad de caja
	Empresas ley 590		Promoción y oferta
	Personas a cargo		Fondos sociales – Foniñez
Personas a cargo beneficiarias	Jóvenes en Jornada escolar complementaria		
Personas a cargo no beneficiarias	Gastos de administración		
Subsidios	Cuota monetaria	Fondos sociales - Fosfec	Postulados
	Subsidio en especie		Subsidios asignados
	Subsidio en servicios		Subsidios pagados
Administración	Saldo para obras		Vacantes
	Remanente		Vinculados

	Excedentes del 55%		Gastos de administración
	Traslado SSF		Gastos de capacitación
	Gastos de administración		Gastos de agencia de gestión
Servicios sociales y no subsidiados	Usuario	Fondos sociales - Ley 115	Gastos de sistema de información
	Uso		Matriculados
	Participante		Graduados
	Población adscrita		Subsidio en pensión
	Transacciones		Subsidio en matrícula
	Número de crédito asignado		Subsidio en bibliotecas
	Número de reserva de hotel / centro recreativo		

El glosario de negocio es el punto de partida y se convierte en el referente y herramienta de consulta; incluso, ayuda a entender el tipo de organización objeto de estudio.

4.5.2 Disciplinas principales

Corresponde en sí a los componentes conceptuales y técnicos que permiten desarrollar y evolucionar la iniciativa, partiendo de la gestión de calidad, datos maestros, metadatos, adicionando actividades de arquitectura, seguridad y modelado en torno a los datos.

La implementación de cada una de las disciplinas, se convierte en un programa o proyecto; por este motivo es importante trabajarlas de manera independiente como entregables pero articuladas como un todo dentro de la iniciativa de gobierno.

En atención a la amplitud (alcance y cronograma) que puede llegar a ser cada disciplina planteada, en este documento únicamente se abordarán los conceptos, desde una mirada muy estratégica, que permitan abarcar los elementos necesarios para iniciar la implementación de la estrategia.

En la elaboración del glosario de negocio se pueden identificar algunas oportunidades de mejora que se traducen en programas y proyectos de gobierno y que son el insumo primordial para la ejecución de las áreas de conocimiento.

La puesta en marcha del cronograma para las disciplinas identificadas, dependerá del sentido de urgencia identificado en el comité de GD, para lo cual, es necesario aplicar los criterios de la matriz de priorización (tabla 5), y de esta manera lograr un proceso organizado y articulado con los objetivos de la gestión de datos.

De igual forma, para apoyar a la organización de las iniciativas, se proponen unos contenidos mínimos para garantizar la adecuada implementación:

- Determinación del nivel de madurez³⁰: en el capítulo 3.1 se realizó un análisis detallado de la metodología, se trataron diferentes componentes y dimensiones concernientes a gobierno, calidad, seguridad, arquitectura, datos maestros, metadatos, etc.; por lo tanto, dicha información se convierte en el insumo adecuado para el despliegue de cada área de conocimiento.

Tener presente que, a nivel general de acuerdo a las dimensiones del CMMI la organización está evaluada dentro del nivel 2; sin embargo, puede existir un componente que se encuentre más o menos desarrollado, de acuerdo al proceso evolutivo de la CCF.

- Misión, visión y objetivos: si bien este contenido ya se definió en las políticas generales del gobierno de datos, es recomendable declarar siempre un propósito por el que se está realizando y adelantando; los inductores o impulsores son los que clarifican el norte y sobre todo limita las expectativas de los clientes finales.

Es posible también incluir el impacto identificado dentro de la matriz de priorización del comité de gobierno de datos; éste se convierte en un motor para acelerar la ejecución.

- Determinar roles y responsabilidades: las áreas involucradas, que se espera que realicen y, sobre todo qué deben realizar en cada momento del proceso, se identifican los proveedores y consumidores que motivan la implementación de la iniciativa.
- Cronograma: detalle de lista de actividades, asignando el tiempo y los responsables, para este ejercicio se plantean semanas por las actividades propuestas, pero es a manera de ejemplo, ya que esto requiere un análisis más profundo desde un punto de vista técnico, para medir un tiempo exacto que toma cada entregable.
- Despliegue: aplicación como tal de las actividades³¹ propuestas, con el ánimo de realizar seguimiento a la ejecución se idean indicadores que midan el avance,

³⁰ Ver anexo 1 con listado de preguntas por cada área de conocimiento.

³¹ La DAMA (2017), ofrece dentro del diagrama de contexto, una detalle de las actividades a desarrollar, las cuales son un referente para la aplicación de acuerdo a las necesidades de la corporación.

por ejemplo, actividades cumplidas / actividades propuestas, encuestas de percepción con usuarios interesados, etc.

4.5.2.1 Definición de estrategia de calidad de datos

Proceso mediante el cual se asegura el ciclo de vida de los datos definidos. Es integral del modelo de gobierno con responsables y actividades establecidas, con el objetivo de garantizar la gestión y control de los datos desde la producción hasta el consumo.

A nivel conceptual son muchos los enunciados para definir la calidad, como punto de partida se puede afirmar que “un dato es de calidad cuando es válido para el propósito para el que un usuario de esos datos quiere utilizarlo” (Batini, C. & M. Scannapieco, 2006).

En las definiciones de Soares (2010), igualmente afirma que se trata de tener datos que sean adecuados para el propósito acordado, esto quiere decir que el criterio de calidad depende del uso de los datos y su respectivo análisis.

La DAMA (2017) define la CD como “el grado en que los datos son precisos, completos, oportunos, consistentes con todos los requisitos y reglas de negocio, relevantes para el uso determinado”, y va mucho más allá para definir también la calidad de la información como “el grado en el que la información cumple consistentemente los requisitos y expectativas de los trabajadores del conocimiento en el desempeño de sus trabajos”; destacando para este caso como consumidores o clientes finales a los usuarios que usan la información para producir conocimiento y sabiduría. Igualmente se listan y conceptualizan las dimensiones de calidad³².

Los estándares ISO realizan a través de la familia 25000, la definición de los conceptos asociados, los cuales son referente para empezar a construir una estrategia de CD.

La ISO/IEC/25012 detalla un modelo para aquellos datos estructurados que se encuentran en un sistema informático (Calabrese, Esponda, Pasini & Otros, 2019). La ISO entiende la calidad como el grado que los datos satisfacen los requisitos de la organización y muestra 15 características agrupadas así:

- **Calidad de datos inherente:** hace referencia al potencial intrínseco cuando los datos se usan con un propósito determinado (1. exactitud, 2. completitud, 3. consistencia, 4. credibilidad, 5. actualidad)

³² Son ejemplos de dimensiones: exactitud, completitud, consistencia, actualización, precisión, privacidad, razonabilidad, integridad, oportunidad, unicidad, validez, accesibilidad (DAMA, 2017).

- **Calidad de datos dependiente del sistema:** hace referencia cuando la calidad es alcanzada por las actividades de un programa informático establecido para este fin (6. accesibilidad, 7. conformidad, 8. confidencialidad, 9. eficiencia, 10. precisión, 11. trazabilidad, 12. comprensión, 13. disponibilidad, 14. portabilidad, 15. recuperabilidad)

Desarrollo del caso práctico:

A continuación, se describen todos los elementos y pasos para plantear la estrategia mencionada:

- Nivel de madurez: de acuerdo a la aplicación del CMMI para determinar el punto de partida, se indagaron componentes de estrategia de CD, perfilamiento de datos, evolución de la calidad y limpieza de los datos. La conclusión general afirma que, si bien existe y se aplican algunas reglas de calidad, el proceso no está formalizado y resulta imperativo su implementación.
- Inductores: garantizar que los datos producidos por la CCF, sean adecuados para el cumplimiento regulatorio y estratégico, aplicando las dimensiones de exactitud, integridad, consistencia; es decir, aplicar la exploración y explotación de los datos sin ningún tipo de sesgo asociado a la CD.
- Misión: generar actividades y programas de CD para gestionar, reducir y eliminar los defectos identificados.
- Visión: ser soporte constante que garantiza la correcta explotación y exploración de datos.
- Políticas:
 - La CD es un proceso que incluye a todos los niveles y está incluido dentro de todas las iniciativas propuestas.
 - Las acciones deben enfocarse hacia garantizar una tendencia a la baja en problemas asociados a una mala calidad.
 - La estrategia de CD proviene desde el comité de GD con el insumo proporcionado por el negocio.
 - Se deben establecer los parámetros y puntos de control para lograr la calidad preventiva.
 - La calidad correctiva, se gestionará desde procesos automáticos para reducir al máximo la intervención manual.
- Ejemplo para la entidad -afiliados-: como se mencionó en apartados anteriores, el punto de partida será el glosario de negocio, el cual indica las entidades que

se pueden priorizar. En la siguiente tabla se muestran algunos atributos asignados para la entidad afiliados:

Tabla 11 Calidad de datos: atributos, dimensiones, reglas, roles y responsabilidades

Entidad	Atributo	Dimensión de calidad	Reglas	Rol / Responsabilidad – corporativo	Rol / Responsabilidad – negocio
Afiliado	Business partner	Exactitud	Registro único, sin duplicados	Rol: analista de calidad de datos Responsabilidad: identificar y notificar excepciones a las reglas planteadas.	Rol: analista de calidad de datos, jefe de área, custodio y propietario del dato Responsabilidad: analizar, corregir y justificar las excepciones presentadas.
	Tipo y número de documento	Exactitud	Mayores de 18 años con tipo de documento CC Longitud de CC con máximo 10 dígitos		
	Género	Integridad	No puede ser vacío		
	Fecha de nacimiento	Exactitud	Superior a 1920 Menor a 1920 se valida con Supervivencia		
	Municipio de residencia	Exactitud	Únicamente Cundinamarca, Boyacá, Tolima		
	Salario y aporte pagado	Consistencia	Menor a \$80 mm		
	Categoría	Consistencia	De acuerdo a Decreto 827 (A-B-C)		
	Segmento poblacional	Consistencia	Alto – básico – medio – joven – independiente - pensionado		
	Dirección y correo electrónico	Integridad	No puede ser vacío		

Nótese que para los atributos³³ seleccionados se asigna una dimensión de calidad y se define la o las reglas a aplicar sobre esa entidad. Cada una de las reglas que se están aplicando apuntan a garantizar una dimensión de calidad. Importante destacar que desde las políticas o mandatos del comité de GD se aprueba el uso y priorización de estos atributos.

En la tabla 12 se realiza una aproximación que sirve de referencia para analizar las dimensiones de calidad con los criterios de negocio aplicados, se incluye también el componente de perfilamiento de datos³⁴, el cual es definido como el proceso de descubrir e investigar problemas de calidad, analizando las fuentes de datos y su origen, para generar estadísticas sobre los mismos:

Tabla 12 Calidad de datos: dimensiones y criterios de negocio

Dimensión de calidad	Criterio de negocio	Perfilamiento de datos
Exactitud	Registros únicos, existentes en la base de datos	Cantidad de registros con esta condición / cantidad total de registros
Integridad	Campos sin vacíos, sin caracteres especiales	Promedio de registros mensuales con esta condición durante los últimos 5 años
Consistencia	De acuerdo a la reglamentación vigente, a los criterios de análisis	Cantidad de registros con esta condición / cantidad total de registros

³³ El atributo se define como la característica de algo que es esencial para su existencia, para este caso se toman como ejemplo algunos conceptos, pero los atributos corresponden y son establecidos de acuerdo a cada entidad identificada.

³⁴ Definición de Gartner, disponible en <https://n9.cl/su8c9>

Actualización	Disponibles en el momento del consumo	-
Precisión	Correspondencia con la realidad de la CCF	-
Privacidad	Garantizar que únicamente los roles autorizados tengan acceso a datos privados	-
Razonabilidad	Coherencia de acuerdo a la operación, transacciones no superiores a cantidad de afiliados, ecuación patrimonial, usos mayores a las personas y participantes.	Promedio de registros mensuales con esta condición durante los últimos 5 años
Oportunidad	Tiempos de alimentación del repositorio de datos.	-
Unicidad	Entidad de datos única, dentro del conjunto definido.	Cantidad de registros con esta condición / cantidad total de registros
Integridad referencial	Trabajadores de acuerdo a las empresas, personas a cargo de acuerdo a los trabajadores.	Promedio de registros mensuales con esta condición durante los últimos 5 años
Validez	Uso de los formatos establecidos, con los atributos necesarios para su análisis	

Para este ejemplo, el rol que identifica la falla de calidad es el analista corporativo, quién se encarga de realizar la notificación al custodio y propietario del dato y así proceder de conformidad, bien sean aplicando el cambio en el sistema fuente o justificando la novedad.

Ahora bien, como se muestra en la siguiente tabla, para cada regla de calidad se determina un indicador y una métrica, asociando el riesgo derivado de una mala calidad:

Tabla 13 Calidad de datos: atributos, riesgos, indicador, métrica y meta

Atributo	Riesgo asociado	Indicador	Métrica	Meta ³⁵
Business partner	Datos incompletos o inconsistentes	Cantidad de defectos / Cantidad de registros correctos	Porcentaje (0-5%)	0%
Tipo y número de documento	Errores en reportes legales	Cantidad de defectos / Cantidad de registros correctos	Porcentaje (0-5%)	0%
Género	Poco conocimiento del cliente	Cantidad de registros vacíos	Valor	0
Fecha de nacimiento	Inoportunidad en las dinámicas comerciales	Cantidad de registros por fuera del rango	Valor	0
Municipio de residencia		Cantidad de municipios por excepción	Valor	0
Salario y aporte pagado	Inoportunidad en las dinámicas comerciales	Cantidad de registros por fuera del rango	Valor	0
Categoría	Inadecuada prestación de servicios	Cantidad de defectos / Cantidad de registros correctos	Porcentaje (0-5%)	0%
Segmento poblacional	Estrategias no focalizadas	Cantidad de registros por fuera del rango	Valor	0
Dirección y correo electrónico	Baja penetración en estrategia de contacto	Cantidad de registros vacíos	Valor	0

³⁵ Para este caso se propone una meta de cero, ya que el objetivo es no tener ningún incidente de calidad, sin embargo, el comité del GD deberá determinar los umbrales mínimos aceptables de defectos por calidad, dependiendo del volumen y de los factores críticos de los atributos analizados.

Como se mencionó en el capítulo anterior, se proponen entidades y sub-entidades en las CCF, sobre las cuales se aplican las reglas, indicadores, roles y responsabilidades.

- Cronograma: para lograr el despliegue de la estrategia de CD, es importante realizar la segmentación de las reglas de calidad, de tal manera que se prioricen de acuerdo al impacto (reputación, legal, estratégico) y de acuerdo al volumen de hallazgos (cambios que pueden ser inmediatos o que pueden tardar tiempo cuando se requiera, por ejemplo, comprar información a alguna entidad del estado). Separando en todo caso, las actividades de análisis de los datos, de las actividades de limpieza como tal.

Se propone el siguiente derrotero para validación y análisis (ver tabla 13), donde adicional al tiempo por cada entidad se incluye el indicador de seguimiento, y así lograr el cumplimiento de los objetivos o medir las brechas de desviación con el plan inicial.

Estas actividades se pueden asociar a una calidad de tipo preventivo (el mejor camino para gestionar la CD es la prevención e incluye actividades como puntos de control en sistemas transaccionales, entrenamiento y capacitación, focalización de reglas, validación de procesos, entre otros):

Tabla 14 Planeación estrategia de calidad de datos

Entidades	Definición de reglas	Definición de indicadores	Roles y responsabilidades	Seguimiento
Afiliados	1 semana	2 semanas	1.5 semanas	Actividad ejecutada / actividad planeada
Subsidios	1 semana	2 semanas	1.5 semanas	
Administración	1 semana	2 semanas	1.5 semanas	
Servicios sociales y no subsidiados	1 semana	2 semanas	1.5 semanas	
Fondos sociales – Fovis	1 semana	2 semanas	1.5 semanas	
Fondos sociales – Fosfec	1 semana	2 semanas	1.5 semanas	
Fondos sociales – Foniñez	1 semana	2 semanas	1.5 semanas	
Fondos sociales – Ley 115	1 semana	2 semanas	1.5 semanas	

Y se propone el siguiente derrotero para el caso de la limpieza de datos (ver tabla 14), con un par de propuestas; las cuales se asociación a una calidad de tipo correctivo (estas actividades ya se presentan cuando han ocurrido los problemas, se pueden utilizar herramientas manuales o automáticas que permitan solventar las situaciones presentadas):

Tabla 15 Planeación limpieza de datos

Estrategia	Propuesta	Riesgo	Volumen	Tiempo	Responsable	Seguimiento
Calidad de datos	Completar registros vacíos en género	Medio	12.000 casos	3 meses	Sección administración de afiliados	Cantidad de registros corregidos
	Validar tipo y número de identificación	Alto	400.000 casos	6 meses	Sección administración de afiliados	

La priorización se realiza de acuerdo al riesgo existente, se valoran en volumen de registros a validar, aplicando un tiempo estimado para realizar dicha actividad y un responsable encargado; al igual que en la fase de validación y análisis, es fundamental incluir el seguimiento al cumplimiento o incumplimiento de las tareas propuestas, en este caso, por ejemplo, la cantidad de registros corregidos, con esto se valida el avance del cronograma y se pueden tomar medidas correctivas en un tiempo determinado (Tabla 15).

Vale recordar que la supervisión está soportada en el líder del GD y, por lo tanto, hay una gestión de seguimiento desde el comité de GD; lo cual indica una profunda integración con las demás iniciativas a que haya lugar y vigilancia en la adecuada aplicación del cumplimiento de las actividades planteadas.

4.5.2.2 Definición de datos maestros y referenciales

La identificación de los datos maestros (DM) y referenciales constituye un punto fundamental y crucial dentro del alcance del espectro del modelo de GD, el requerimiento de tener los mismos datos e información dentro de todos los niveles de la organización, cobra especial relevancia; hay que anotar que, dentro del modelo de una organización actual, los datos son el factor diferenciador.

La definición de la DAMA (2017) para DM reza que son aquellos relativos a las entidades de negocio que dan un contexto a las transacciones de negocio generadas (disponibles y precisos para las entidades analizadas), se refiere a clientes, ciudadanos, pacientes, vendedores, proveedores, socios, competidores, productos, lugares, datos financieros, etc.

Desde el punto de vista técnico, con el ánimo de empujar la estrategia de DM, se propone la creación de una bodega / repositorio único de datos y, tener con las definiciones claras una única fuente de verdad. Aclarando que sólo se constituye en la herramienta para gestionar la estrategia, no es el fin como tal y no quiere decir que con dicha actividad ya se finalice la estructuración de los DM.

Dentro de esta estrategia también resulta importante la definición de roles y responsabilidades, sin embargo, es un tema que se encuentra inherente dentro del

modelo operativo, y ya debería estar cubierto en dicha iniciativa. No obstante, no es redundante incluirlo nuevamente dentro de esta estructura.

Desarrollo del caso práctico:

A continuación, se describen todos los elementos y pasos para plantear la estrategia mencionada:

- Nivel de madurez: si bien dentro de la metodología de la medición del estado actual no se incluyó esta capacidad, su puesta en marcha es indispensable para soportar las actividades de CD y por lo tanto del GD en su conjunto.
- Inductores: lograr articulación entre las iniciativas y el GD, con el despliegue de la estrategia de DM se busca apalancar la CD y tener un Golden Record³⁶ acordado, definido y establecido; el cual se obtiene para cada entidad relevante, a partir de la cual se derivan los procesos de análisis más importantes para las CCF.
- Misión: integrar las diferentes capacidades de la organización para establecer datos maestros y de referencia, para la gestión estratégica de los datos.
- Visión: crear una cultura uniforme de datos maestros y de referencia, sobre la cual todos los clientes de la organización vean a la estrategia de DM como el referente y fuente de verdad.
- Políticas:
 - La planeación y asignación de roles y responsabilidades se realizará desde el comité de GD.
 - En el negocio, se mantiene y evoluciona la estrategia de DM.
 - Aplicar estándares y buenas prácticas que permitan la adhesión de los DM en todas las áreas de la CCF.
 - Será responsabilidad del custodio y propietario de los datos, la correcta aplicación de la estrategia de DM.
 - Los DM serán insumo para todo programa y proyecto asociado a la gestión de datos.
- Ejemplo: continuando con la explotación y uso de la herramienta del glosario de negocio, se toman como referencia las entidades ya identificadas:

³⁶ Cualidad que define la versión única de las entidades, es decir, el dato e información correctos.

En la siguiente tabla se muestran las entidades, con la identificación de dato maestro:

Tabla 16 Identificación de datos maestros y referenciales

Estrategia	Entidades	Dato maestro	Atributo	Dato referencial	Dato crítico	Relación	Sistema fuente ³⁷
Datos maestros y referenciales	Afiliados	Afiliado – persona Afiliado – empresa	Fecha de afiliación	Tipos de documento Estado de afiliado	X	Con procesos de: Reportes legales Estrategia de segmentación de clientes Proyectos de innovación Contabilidad Mercadeo Productos	Cada uno de los sistemas transaccionales usado por la CCF, y los sistemas donde se encuentra el repositorio de datos.
	Subsidios	Beneficiario	Fecha de asignación	Tipos de documento	X		
	Administración	Cuenta contable	Tipo de cuenta	Agrupación cuenta	X		
	Servicios sociales y no subsidiados	Cliente alojado - hoteles Cliente atendido - supermercados Cliente atendido - salud IPS	Fecha de transacción		X		
	Fondos sociales – Fovis	Postulado	Fecha de postulación	Estado del postulado - Fovis	X		
	Fondos sociales – Fosfec	Postulado	Fecha de postulación	Estado del postulado -Fosfec	X		
	Fondos sociales – Fonñez	Niño atendido Joven beneficiario	Modalidad	Tipos de documento	X		
	Fondos sociales – Ley 115	Estudiante matriculado			X		

Ahora, en la siguiente tabla se muestran las definiciones de los datos referenciales.

Tabla 17 Definición de datos referenciales

Dato referencial	Descripción
Tipos de documento	1. Cédula de ciudadanía 2. Tarjeta de identidad 3. Registro civil de nacimiento 4. Cédula de extranjería 5. Permiso especial de permanencia
Estado de afiliado – persona	1. Activo 2. Retirado 3. Retirado con derecho a subsidio 4. Retirado en período de gracia
Estado de afiliado – empresa	1. Activa 2. Retirada 3. Expulsada
Agrupación cuenta	1. Activo 2. Pasivo 3. Patrimonio 4. Ingresos 5. Gastos 6. Costos
Estado del postulado – Fovis	1. Asignado 2. Pagado 3. Vencido 4. Restituido 5. Renunciado 6. Restablecido
Estado del postulado -Fosfec	1. Asignado 2. Girado 3. Reactivado 4. Reintegrado

³⁷ Hace referencia a los sistemas tanto productores como consumidores de datos.

Para lograr una correcta ejecución de este componente se sugieren los siguientes pasos:

- a. **Tipificación de entidades de datos:** de acuerdo al inventario de datos que posee la empresa, se realiza la identificación de los dominios que serán relevantes para la operación de la gestión de los datos, se recomienda realizar ordenamiento de mayor a menor relevancia (número de documento, dirección, teléfono, compras, transacciones, cuenta contable, centro de costo, etc.), de acuerdo a la competencia estratégica de datos definida³⁸.
 - b. **Asignación de criterio de “dato crítico”:** las entidades de datos definidos se marcan como críticos o no críticos, de acuerdo a la validación del evaluador, ya que dependerá del análisis y los objetivos que se estén realizando. Entre algunos criterios para esta actividad, podrían ser: cantidad de áreas que usan el dato, necesidad de contacto con el cliente, impactos en reportes legales, entre otros.
 - c. **Identificación de relaciones:** las dependencias e interacciones de los datos maestros son identificadas para medir el impacto que se pueda generar ante cualquier ajuste y/o modificación. El linaje y la trazabilidad que se pueda adelantar sobre los datos maestros ayudarán en la tipificación del ciclo de vida de los datos.
 - d. **Caracterización de sistemas fuente:** mapeo de los datos desde los sistemas transaccionales de la empresa, para asegurar la única fuente de verdad desde el repositorio de datos establecido.
- Cronograma: el objetivo es realizar la planeación a alto nivel de las actividades a ejecutar y, lograr aplicar los objetivos de la estrategia (Tabla 18).

Tabla 18 Planeación estrategia datos maestros y referenciales

Estrategia	Entidades	Definición de dato maestro	Definición de dato referencial	Asignación de criterios, relaciones y sistema fuente	Responsable	Seguimiento
Datos maestros y referenciales	Afiliados	1.5 semanas		2 semanas	Gerencia / área asociada al dato maestro (custodio o propietario)	Actividad ejecutada / Actividad planeada
	Subsidios	1.5 semanas		2 semanas		
	Administración	1.5 semanas		2 semanas		
	Servicios sociales y no subsidiados	1.5 semanas		2 semanas		
	Fondos sociales – Fovis	1.5 semanas		2 semanas		
	Fondos sociales – Fosfec	1.5 semanas		2 semanas		

³⁸ La estrategia de datos puede estar definida de acuerdo a los objetivos iniciales del gobierno de datos: calidad, contactabilidad de clientes, identificación de productos, mejora en procesos, etc.

	Fondos sociales – Foniñez	1.5 semanas	2 semanas		
	Fondos sociales – Ley 115	1.5 semanas	2 semanas		

4.5.2.3 Definición de metadatos

Teniendo en cuenta que la finalidad de los metadatos (MD) es la de describir³⁹ y dar información sobre los datos, es decir, de ser un catálogo del inventario de activos que posee la corporación; aparece la exigencia de adoptar dicha estrategia dentro del espectro del gobierno y la gestión de datos.

La estrategia y por lo tanto la definición de los metadatos, está directamente relacionado con el glosario de negocio y con los datos maestros, se articulan desde el entendimiento y conocimiento de los datos de la organización para dar paso a una gestión eficiente de los mismos.

Por la diversidad de operaciones de las CCF y los servicios que están habilitadas para prestar⁴⁰, la acogida de herramientas que permitan gobernar y estandarizar, serán siempre bienvenidas.

La recomendación es aplicar la estrategia de metadatos en dos niveles:

1. Sobre el repositorio único de datos, ya que existe una mayor estandarización que en los sistemas transaccionales, lo cual no quiere decir que no se pueda aplicar sobre otro tipo de solución.
2. Sobre los sistemas transaccionales, ya que es ahí donde de donde se van a obtener los datos para la bodega, el objetivo es conocer los datos migrados con sus características y reglas de negocio necesarios

Desarrollo del caso práctico:

A continuación, se describen todos los elementos y pasos para plantear la estrategia mencionada:

- Nivel de madurez: en la medición del estado actual y general que se hizo para este caso de uso, se abordó el componente de gestión de metadatos; arrojando comentarios sobre la no existencia de un programa de MD y, por lo tanto, no es un quehacer dentro de las actividades de los roles actuales.

³⁹ <https://rb.gy/g3eib1>

⁴⁰ Como se ha mencionado previamente, esta situación deriva en tener diferentes sistemas transaccionales para soportar la operación de las unidades de servicio.

- Inductores: propiamente la carencia de MD lleva al planteamiento de dicha metodología, para apalancar la generación de valor de los datos, aplicando un lenguaje único y común, que permita la reducción de los costos por redundancia de datos e información; la aplicación de esta estrategia también soporta las actividades de CD.
- Misión: aplicar los estándares definidos en la estrategia de MD para generar un directorio estable, unificado y único, que cumpla con los estándares requeridos por los usuarios finales.
- Visión: ser un referente para los clientes de los datos, en torno a la construcción de una metodología que se ajuste a la gestión estratégica de los metadatos.
- Políticas:
 - Deberá existir y se desarrollará una estrategia de metadatos en la CCF.
 - Los MD serán herramienta de consulta para toda iniciativa de gestión de datos.
 - La gestión de la demanda de nuevos reportes, análisis y usos de los datos, estará alineado con la estrategia de MD.
- Ejemplo: en la siguiente tabla se muestra la caracterización de los elementos a incluir dentro de los MD para el dato maestro de afiliado – persona, desde el tipo de negocio⁴¹.

Tabla 19 Identificación de metadatos

Estrategia	Entidades		
Metadatos	Afiliados	Dato maestro	Afiliado – trabajador dependiente
		Dato referencial	Tipos de documento Estado de afiliado
	Definición	Los trabajadores de carácter permanente al servicio de los empleadores previstos en los artículos 7o. y 72 de la ley 21 de 1982, desde el momento de su vinculación y hasta la terminación de la misma.	
	Política de calidad (reglas)	Mayores de edad con tipo de documento cédula de ciudadanía Menores de edad con permiso para trabajar en el sector formal Trabajadores correspondientes a empresas activas	
	Política de seguridad (componente)	Público	
	Bodega y objeto	BW afiliaciones y BW-SAP-AFI-01	
	Formato	Numérico	
	Longitud	7 caracteres	

⁴¹ De acuerdo a la DAMA (2017), adicional negocio, existe el tipo de metadato técnico (acerca del sistema fuente, de la bodega de almacenamiento y sus procesos derivados) y operativo (se refiere a historial de consultas, resultados de auditoría, reportes y acceso, entre otros).

	Calculado	No
	Estado	Activos Retirados con derecho a subsidio monetario
	Relación con otros datos maestros	Empresas, personas a cargo, sistemas transaccionales
	Actualización	Diaria a media noche, desde sistemas transaccionales a repositorio
	Responsable	Gerencia de Operación de Subsidios
	Impacto ⁴²	Alto
	Modelo de datos	Ver apartado Modelado y Diseño
	Atributos	Business Partner, tipo y número de documento, género, fecha de nacimiento, municipio de residencia, salario y aporte pagado, categoría, segmento poblacional, dirección y correo electrónico
	Procesos dependientes	Reportes legales, Estrategia de segmentación, Dinámicas comerciales, Contacto con el cliente

- Cronograma: para lograr el objetivo de implementar un modelo de metadatos en la organización, se proponen las siguientes actividades; las cuales ya tienen un nivel de madurez alto porque se han trabajado desde el glosario de negocio y la definición de los datos maestros (Tabla 20).

Al igual que en apartados anteriores, se destaca al responsable del negocio como principal movilizador y ejecutor de la iniciativa y un capítulo de seguimiento que con indicador definido permite el monitoreo del cumplimiento de los objetivos trazados para éste propósito.

Tabla 20 Planeación estrategia de metadatos

Estrategia	Entidades	Definición de metadatos	Asignación de criterios	Responsable	Seguimiento
Metadatos	Afiliados	1.5 semanas	2 semanas	Gerencia / área asociada al dato maestro (custodio o propietario)	Actividad ejecutada / Actividad planeada
	Subsidios	1.5 semanas	2 semanas		
	Administración	1.5 semanas	2 semanas		
	Servicios sociales y no subsidiados	1.5 semanas	2 semanas		
	Fondos sociales – Fovis	1.5 semanas	2 semanas		
	Fondos sociales – Fosfec	1.5 semanas	2 semanas		
	Fondos sociales – Foniñez	1.5 semanas	2 semanas		
	Fondos sociales – Ley 115	1.5 semanas	2 semanas		

⁴² Se incluye criterio de impacto, con el ánimo de aplicar el concepto de criticidad del dato para los operaciones de la Corporación (alto, medio, bajo).

4.5.3 Disciplinas de soporte

Incluye la arquitectura, seguridad y privacidad, modelado; elementos que apoyan otras estrategias y que son necesarias dentro del despliegue de los programas de gestión y gobierno de datos.

4.5.3.1 Definición de arquitectura de datos

La estrategia que se definirá en este capítulo está precisamente ligada a la definición de los requisitos de negocio, la integración y el control de los activos, que es el concepto de arquitectura; en la cual se incluye definiciones, estructuras, reglas y documentación robusta (DAMA, 2017).

De acuerdo con Marchildon (2018), es el diseño de todas las aplicaciones y sistemas, que hace que los datos estén disponibles para todos los consumidores. Apoya las necesidades de datos de la empresa y permite la estandarización e integración.

Incluye tres categorías: modelo de datos (ver capítulo 4.5.3.3), cadena de valor de la información (ver capítulo 4.3) y la arquitectura, integración, almacenamiento de las bases de datos. Igualmente, la organización detallada de los componentes hace parte de ésta área de conocimiento.

Desarrollo del caso práctico.

A continuación, se describen todos los elementos y pasos para plantear la estrategia mencionada:

- Nivel de madurez: dentro del diagnóstico realizado en la etapa previa, se indagó sobre el capítulo de plataforma y arquitectura, el cual arroja conclusiones como: no hay un enfoque hacia los activos, no se percibe un conjunto de tecnología de datos aprobado, no son claros los niveles de autoridad y no se conoce sobre los modelos de datos; sin embargo, la implementación de la arquitectura de los datos para la gestión del negocio es inminente.
- Inductores: desarrollar un enfoque sistémico de la gestión de datos y asegurar los temas de la agenda del GD, es decir, gestionar la demanda con el consumo; traducir las necesidades de negocio en requerimientos de datos.
- Misión: proporcionar herramientas que permitan entender, identificar, ordenar y gestionar los recursos asociados a los datos de la Corporación.
- Visión: ser un modelo de arquitectura alineado con las políticas del gobierno de datos, para asegurar la continuidad del negocio.

- Ejemplo: en la tabla 21 se puede visualizar la arquitectura empresarial a un alto nivel para la entidad de afiliados.

Tabla 21 Arquitectura empresarial para la entidad afiliados

Matriz	Qué	Cómo	Dónde	Quién	Cuándo	Por qué
Contextual	Reporte de afiliados	A demanda	n/a	Líder de gobierno de datos	A demanda	Análisis interno
Conceptual	Contiene listado de afiliados activos a la CCF	Solicitar, extraer y enviar	n/a	Coordinador de modelos	Diario	Análisis externo
Lógico	n/a	Diagrama de flujo	Sistema transaccional	Jefe de Afiliaciones		Semanal
Físico	n/a		Repositorio / Bodega	Analista de aplicaciones		
Componente	n/a		BW-SAP-AFI-01			
Operacional	Gerencia de Operación de Subsidios		n/a		Jefe de Afiliaciones	A demanda

Fuente: elaboración propia a partir de Zachman (1980)

En la tabla 22 se conceptualiza la arquitectura de datos para la misma entidad a un alto nivel, la cual incluye la estrategia, requerimiento de negocio, propósito y el modelo basado en flujo de datos.

Tabla 22 Arquitectura de datos para la entidad afiliados

Estrategia	Entidad	Requerimiento de negocio	Propósito	Arquitectura ⁴³
Arquitectura	Afiliados	Obtener reporte semanal de afiliados a la CCF	Seguimiento a la estrategia comercial de atracción de afiliados	

⁴³ Dentro de la metodología de la DAMA (2017) a nivel de arquitectura empresarial de datos, están dos opciones: modelo de datos empresariales y diseño de flujos de datos.

- Cronograma: tomando como referencia nuevamente el listado de entidades ya definidas, se propone el mapa de ruta descrito en la tabla 23, se detallan los requerimientos de negocio como principal inductor, definiendo los dos elementos de arquitectura (empresarial y datos), y como en todas las estrategias la identificación de responsables y el capítulo de seguimiento.

Tabla 23 Planeación estrategia de arquitectura de datos

Estrategia	Entidades	Requerimiento de negocio	Arquitectura empresarial	Arquitectura de datos	Responsable	Seguimiento
Arquitectura de datos	Afiliados	1.5 semanas	2 semanas		Gerencia asociada / Oficina de Tecnología	Actividad ejecutada / Actividad planeada
	Subsidios	1.5 semanas	2 semanas			
	Administración	1.5 semanas	2 semanas			
	Servicios sociales y no subsidiados	1.5 semanas	2 semanas			
	Fondos sociales – Fovis	1.5 semanas	2 semanas			
	Fondos sociales – Fosfec	1.5 semanas	2 semanas			
	Fondos sociales – Foniñez	1.5 semanas	2 semanas			
	Fondos sociales – Ley 115	1.5 semanas	2 semanas			

4.5.3.2 Definición de estrategia de seguridad y privacidad

El componente de seguridad de los datos y de la información cobra especial relevancia, en primer lugar, a nivel externo existen varias regulaciones⁴⁴ que obligan a implementar acciones para la protección y control y, en segundo lugar, a nivel interno el interés por salvaguardar estos activos es más que necesario.

Al convertirse los datos, la información y el conocimiento en el insumo diferenciador de éxito o fracaso empresarial, es menester aplicar todas las herramientas necesarias para garantizar su correcta seguridad y privacidad.

En la definición de la DAMA (2017), se afirma que incluye todas las actividades de planeación y ejecución de las políticas y procedimientos para proveer autenticación, autorización, acceso y auditoría a los datos y los activos de información.

Desarrollo del caso práctico:

A continuación, se describen todos los elementos y pasos para plantear la estrategia mencionada:

⁴⁴ Ley 527 de 1999, Ley 1266 de 2008, Ley 1273 de 2009, Ley 1581 de 2012, entre otras.

- Nivel de madurez: si bien existe un área formal de seguridad, es necesario realizar articulación con el programa de GD como un todo. Las iniciativas de seguridad no son exclusivas del área de tecnología, por el contrario, son inclusivas para todas las áreas de negocio y hacen parte de la operación diaria de los custodios y propietarios de los datos e información.

En este sentido, se propone una serie de actividades que permitan impregnar la cultura de seguridad en todos los niveles.

- Inductores: crear y aplicar las buenas prácticas de un modelo efectivo de seguridad de la información.
- Misión: generar actividades y programas asociados a los activos de información a custodiar, para identificar riesgos y políticas.
- Visión: ser la estrategia con el mayor nivel de desarrollo e innovación dentro del programa de GD, garantizando la continuidad del negocio por medio de buenas prácticas que permiten proteger a los datos e información de amenazas y vulnerabilidades.
- Políticas:
 - Definir e identificar repositorios de custodia de los activos de datos e información.
 - Definir roles y perfiles (antes, durante y después de la vinculación laboral – ciclo de vida de los trabajadores).
 - Centralizar proceso de creación y eliminación de activos de datos e información, con un proceso de gestión de la demanda.
 - Crear protocolos de copias de seguridad.
 - Investigar y aplicar mejores prácticas de seguridad, garantizando los protocolos de antivirus, fuga de información y gestión de amenazas.
 - Establecer matriz de accesos a los sistemas transaccionales.
- Ejemplo: para continuar con la ilustración de la implementación de un modelo de seguridad de la información, se propone seguir el siguiente proceso de alto nivel:

Identificación de activos – clasificación de activos – identificación de riesgos.

Para efectos prácticos, únicamente se realiza el análisis a nivel de entidad, pero el nivel de profundidad se puede determinar por la sub-entidad y para cada uno los atributos, de acuerdo a la criticidad y grado de seguridad⁴⁵.

Identificación de activos: el punto de partida para esta actividad, será nuevamente el glosario de negocio, en la siguiente tabla se incluyen las variables a destacar.

Tabla 24 Seguridad de datos e información: identificación de activos

Entidad	Atributo	Custodio	Propietario	Tipo de ubicación	Usuario final	Tipo de activo	Formato de conservación
Afiliado	Business partner	Coordinador de procesamiento de datos	Administración de afiliados	Repositorio: BW Afiliaciones	Cualquier cargo asociado dentro del modelo de GD.	Dato	Magnético
	Tipo y número de documento					Dato	Magnético
	Género					Dato	Magnético
	Fecha de nacimiento					Dato	Magnético
	Municipio de residencia					Dato	Magnético
	Salario y aporte pagado					Dato	Magnético
	Categoría					Información	Magnético
	Segmento poblacional					Información	Magnético
	Dirección y correo electrónico					Dato	Magnético

Tipo de ubicación: hace referencia al tipo de lugar donde se encuentra alojado, puede ser un repositorio, carpeta, file Server, etc.; se anota específicamente el nombre o dirección como se puede identificar claramente.

Usuario final: son los clientes que usan los datos para los propósitos internos y externos.

Tipo de activo: para este ejercicio, se hace referencia a si es dato o información, ya que los activos de información pueden ser redes, software, personal, procesos, etc.

⁴⁵ Existen datos públicos (contenidos en registros públicos, censales, profesionales), privados (gustos, preferencias, relaciones), semi-públicos (financieros, crediticios), sensibles (salud, ideología, creencias, vida sexual, biométricos), protección especial (menores de edad, condición de discapacidad). Ley 1581 de 2012, protección de datos personales.

Formato de conservación: aplicaría el medio magnético y físico en casos muy específicos.

Clasificación de activos (Tabla 25): se incluye el tipo (estructurado, semiestructurado y no estructurado); en la dimensión de seguridad se valora el grado de confidencialidad (público, privado, uso interno, reservado), integridad (alto, bajo, medio) y disponibilidad (alto, bajo, medio).

Tabla 25 Clasificación de activos por tipo y dimensión

Entidad	Atributo	Clasificación por tipo	Clasificación por dimensión		
			Confidencialidad	Integridad	Disponibilidad
Afiliado	Business partner	Estructurado	Uso interno	Medio	Medio
	Tipo y número de documento		Público	Medio	Medio
	Género		Público	Medio	Medio
	Fecha de nacimiento		Público	Medio	Medio
	Municipio de residencia		Privado	Alto	Alto
	Salario y aporte pagado		Privado	Alto	Alto
	Categoría		Público	Medio	Medio
	Segmento poblacional		Uso interno	Medio	Medio
	Dirección y correo electrónico		Privado	Medio	Medio

Identificación de riesgos: para este propósito se exponen los siguientes elementos:

- a. **Caracterización:** breve descripción del suceso o evento (qué, cómo, cuándo), atribuido a cuáles fuentes (personas, procesos, tecnología, infraestructura, externos), determinado por qué factores (reputación, legal, operativo, contagio).
 - b. **Valoración:** medir el impacto (catastrófico, mayor, moderado, menor, insignificante) del suceso o evento, con cierta probabilidad (casi cierta, probable, posible, improbable, rara), y con una implicación traducida en severidad (extrema, alta, baja, media)
 - c. **Control:** identificación de los puntos principales a vigilar, determinado uno o varios responsables, con un grado de automatización, detallando la documentación necesaria, evaluando la complejidad y guardando registros históricos.
- Cronograma (Tabla 26): atendiendo las actividades planteadas para la ejecución de la estrategia de seguridad de datos e información, se plantea un ejemplo de la planeación de acciones para las entidades identificadas, igualmente la inclusión del seguimiento.

Tabla 26 Planeación estrategia de seguridad de datos e información

Estrategia	Entidades	Identificación de activos	Clasificación de activos	Identificación de riesgos			Seguimiento
				Caracterización	Valoración	Control	
Seguridad de datos e información	Afiliados	1 semana	1 semana	1.5 semanas			Actividades cumplidas / actividades propuestas
	Subsidios	1 semana	1 semanas	1.5 semanas			
	Administración	1 semana	1 semanas	1.5 semanas			
	Servicios sociales y no subsidiados	1 semana	1 semanas	1.5 semanas			
	Fondos sociales – Fovis	1 semana	1.5 semanas	2 semanas			
	Fondos sociales – Fosfec	1 semana	1.5 semanas	2 semanas			
	Fondos sociales – Foniñez	1 semana	1.5 semanas	2 semanas			
	Fondos sociales – Ley 115	1 semana	1.5 semanas	2 semanas			

4.5.3.3 Definición de estrategia de modelado⁴⁶ y diseño

En este capítulo se abordará la estrategia de representar y comunicar los requerimientos de datos, es un proceso de entendimiento de los datos de la organización, de captura de conocimiento acerca de los activos definidos. El modelado directamente relacionado con el componente de metadatos, ya que el primero es una forma del segundo, cuyo objetivo es lograr un vocabulario común.

Es una herramienta que permitirá apoyar la implementación de estándares, definir alcance de proyectos y así garantizar la retención del conocimiento explícito de la organización.

Al igual que en otros capítulos, el punto de partida es el glosario de negocio ya definido en el apartado 4.5.1.4, cobrando especial relevancia el concepto de entidad.

Desarrollo del caso práctico.

A continuación, se describen todos los pasos para plantear la estrategia mencionada:

- Nivel de madurez: en el ejercicio de diagnóstico se abordó en la pregunta 72 y 73 lo referente a modelado de datos, sin embargo, el nivel operativo transversal está en desacuerdo y en nivel de negocio está acuerdo, lo que lleva a concluir que no

⁴⁶ En DAMA (2017) se define modelo como la representación de algo que existe o algo para ser hecho.

hay una definición clara y visible para este frente; de ahí el requerimiento de su implementación.

- Inductores: esta estrategia nace como necesidad de avanzar en la implementación de un lenguaje común de los datos; para establecer una herramienta de consulta obligada.
- Misión: desarrollar estrategia de modelado, para estandarizar los procesos asociados a la gestión de datos.
- Visión: crear una metodología embebida dentro de la estrategia de gestión y gobierno de datos, que nutra todas las iniciativas asociadas.
- Ejemplo: en la tabla 27 se presenta la entidad afiliado, identificado atributos y dominios⁴⁷ y de manera complementaria se presenta el esquema⁴⁸ (ilustración 23), que es la herramienta usada para pintar los datos con sus notaciones.

Tabla 27 Definición de entidad, atributo y dominio

Entidad	Atributo	Dominio	Esquema
Afiliado	Activo (A) Retirado (R) Expulsado (E)	Estado	Dimensional
	Fecha de nacimiento (N)	Fecha	
	Fecha de afiliación (A)		
	Salario y aporte pagado	Estratégicos	
	Categoría		
	Segmento poblacional		
	Dirección y correo electrónico	Geográficos	
	Municipio de residencia		

En la ilustración 23 se materializa el modelo dimensional para la entidad Afiliado, en la cual se destacan los dominios estratégicos, geográficos, de estado y de fecha, de acuerdo a cada atributo identificado.

⁴⁷ Definido como el medio para estandarizar las características de los atributos.

⁴⁸ En DAMA (2017) adicional al esquema dimensional, se presentan los relacionales, orientados a objetos, basados en hechos, basados en tiempo, NoSQL.



Ilustración 23 Modelo dimensional para la entidad afiliado

Fuente: ilustración propia

En la ilustración 24 se visualiza el modelo conceptual para la entidad afiliado, relacionando en el primer nivel la dependencia entre las personas a cargo que no pueden existir sin afiliados y los afiliados no pueden existir sin empresas. En el segundo nivel se muestra la llave relacional que es la identificación de cada entidad, para el caso de afiliados puede haber relación con varias empresas y en el caso de las personas a cargo, la relación es uno a uno.



Ilustración 24 Modelo conceptual para la entidad afiliado

- Cronograma: se proponen las actividades, tiempos, responsables y seguimiento, como se muestra en la tabla 28.

Tabla 28 Estrategia de modelado de datos

Estrategia	Entidades	Atributos y dominios	Esquemas	Responsables	Seguimiento
Modelado de datos	Afiliados	1 semana	1 semana	Gerencia de Tecnología. Gerencia de Operación de Subsidios.	Actividades cumplidas / actividades propuestas
	Subsidios	1 semana	1 semanas		
	Administración	1 semana	1 semanas		
	Servicios sociales y no subsidiados	1 semana	1 semanas		

	Fondos sociales – Fovis	1 semana	1.5 semanas		
	Fondos sociales – Fosfec	1 semana	1.5 semanas		
	Fondos sociales – Fonñez	1 semana	1.5 semanas		
	Fondos sociales – Ley 115	1 semana	1.5 semanas		

4.6 Cultura y cambio organizacional

Dada la importancia de este componente dentro del despliegue de cualquier iniciativa empresarial, se revela como un capítulo independiente para asegurar completamente la asimilación e implementación de una verdadera cultura del dato y la información.

La cultura y cambio dentro de la Corporación es fundamental para garantizar que, los objetivos estratégicos planteados para el programa de GD lleguen a feliz término, garantizando trabajos cohesionados entre los equipos acordes a los planes y proyectos adelantados (Morelos-Gómez, 2014)

Para lograr permear al público objetivo del GD, se plantean los objetivos y las acciones a ejecutar para lograr que rápidamente la organización asuma los datos y la información como parte de la rutina diaria de los roles involucrados.

4.6.1 Objetivos

Establecer una metodología para la asimilación de los datos y la información como parte inherente y fundamental de todos los procesos de la empresa.

Despliegue del plan de comunicación de acuerdo a los públicos objetivos definidos con temáticas específicas.

4.6.2 Acciones

Se proponen las siguientes actividades para soportar la estrategia de cultura organizacional:

- Charlas abiertas a toda la comunidad con expertos del tema de la gestión de datos, gobierno de datos y las disciplinas asociadas, espacios tipo “webinar” que permiten un acercamiento con casos prácticos y reales que sirven como referente y buenas prácticas a aplicar.
- Estructuración de talleres con el semillero de gestión de la información, que conduzca a formar a los posibles movilizadores de información de la empresa y a identificar posibles programas y proyectos.

- Sensibilización de los datos y la información a través de un boletín diario, semanal o mensual en los que se menciona los alcances y los avances del GD. Igualmente incluir los resultados principales de la compañía, donde se visualiza el trabajo y el esfuerzo del equipo.
- Incluir al documento de gobierno de datos como un componente visible de la estrategia de ética y transparencia corporativa.
- Incluir el esquema de gobierno de datos como componente principal de capacitación e inducción al nuevo personal, de esta manera se garantiza que el tema conocido por todos.
- Identificación de movilizadores visibles dentro del programa, que se encargarán de canalizar todas las decisiones del gobierno a los equipos de trabajo asignados.
- Definición de los contenidos temáticos de los movilizadores, del semillero y de la capacitación a implementar.
- Fortalecer los canales de comunicación para cubrir todas las iniciativas que los públicos objetivos, deseen que forme parte de la agenda del comité del gobierno de datos.
- Encuesta de percepción a usuarios definidos, en la cual se pueden abordar los temas de gobierno de datos y los programas, para identificar el grado de adherencia y conocimiento en la organización (ver anexo 2). En el ejemplo se muestra una estructura de preguntas tipo para la estrategia de gobierno de datos.
- Creación del decálogo⁴⁹ del gobierno de datos, donde se relacionan los 10 temas más relevantes, el cual debe ser de conocimiento y estar al alcance de todos. Se convierte en un documento visible de fácil acceso, que constantemente se está comunicando por los medios de la organización.

4.6.3 Indicadores de seguimiento

Con el objetivo de medir el impacto de las iniciativas definidas en el contexto de la cultura organizacional, se proponen los indicadores descritos en la tabla 29.

⁴⁹ Reglas que se consideran básicas para una actividad

Tabla 29 Listado de indicadores medición de cultura organizacional

Consecutivo	Indicador	Objetivo	Frecuencia
I-01	Cantidad de charlas mensuales	Medir el alcance de las charlas realizadas	Mensual
I-02	Cantidad de participantes en el semillero	Medir el alcance de la cantidad de personas impactadas	Trimestral
I-02	Porcentaje de adherencia y adhesión	Monitorear el nivel de conocimiento de las iniciativas implementadas	Semestral
I-03	Cantidad de talleres desarrollados	Monitorear la cantidad de talleres y los impactos en los participantes	Trimestral
I-04	Cantidad de boletines publicados	Medir la cantidad de productos de información generados	Semestral

5 RECOMENDACIONES

Dentro de este capítulo se tratarán a manera de cierre, dos temas cruciales para balancear la iniciativa de gobierno; en primer lugar, una vez definida la estrategia, se traza la carta de navegación que permita ejecutar en un tiempo definido los programas y proyectos necesarios y, en segundo lugar, los indicadores que permitirán realizar la medición, impacto y seguimiento a las actividades propuestas.

5.1 Mapa de ruta

Tomando como referencia a Soares (2010), el éxito para abordar este entregable es partir de la medición del nivel de madurez, identificar las estrategias clave con las personas, procesos y tecnología necesarias y realizar la priorización de acuerdo a los temas claves o estratégicos para la Corporación.

Se plantean una serie de actividades que sean de guía y referencia al decidir emprender la implementación del gobierno de datos. Basados en la experiencia de este caso de uso se propone seguir los pasos visualizados en la tabla 30.

Para cada uno de los procesos, se establece la iniciativa a desarrollar y se detalla la actividad que se adelanta para el cumplimiento de la estrategia; como los pilares del GD son los componentes, se marca el impacto o requerimiento necesario para el éxito.

La hoja de ruta planeada se convierte en la carta de navegación para lograr una adecuada implementación de las iniciativas de los datos.

Tabla 30 Mapa de ruta para gobierno de datos e iniciativas

Proceso	Iniciativa	Actividad	Componente ⁵⁰		1	2	3	4	5	6	7	8
			Pe.	Pr.								
Planeación	Factores de éxito	Definir y formalizar misión, visión, objetivos, alcance e impulsores.	Pe.	X								
			Pr.	X								
			TI									
	Estructura de trabajo	Definir y formalizar estructura organizacional	Pe.	X								
			Pr.	X								
			TI									
		Definir y formalizar comité de gobierno de datos	Pe.	X								
			Pr.	X								
			TI									

⁵⁰ Componentes o pilares fundamentales para el GD, descritos en el apartado 4.1 (Pe = personas, Pr = procesos, TI = tecnología). El tiempo se encuentra expresado en semanas.

			TI																						
		Identificar y aprobar roles y responsabilidad	Pe.	X																					
			Pr.																						
			TI																						
		Detallar y socializar políticas	Pe.	X																					
			Pr.																						
			TI																						
		Estructurar gestión operativa de los datos	Pe.	X																					
			Pr.	X																					
	TI																								
	Cadena de valor y ciclo de vida	Identificar macro-procesos y etapas de los datos.	Pe.																						
			Pr.	X																					
TI																									
Ejecución	Gestión productiva y operativa	Establecer glosario de negocio	Pe.																						
			Pr.	X																					
			TI	X																					
		Establecer y aprobar activos de datos e información	Pe.																						
			Pr.	X																					
			TI	X																					
		Establecer y aprobar formatos y estructuras de metadatos	Pe.																						
			Pr.	X																					
			TI	X																					
		Establecer y aprobar formatos y estructuras de datos maestros	Pe.																						
			Pr.	X																					
			TI	X																					
		Establecer y aprobar metodología y desarrollo de modelado de datos	Pe.																						
			Pr.	X																					
			TI	X																					
		Formalizar proceso de calidad de datos	Pe.	X																					
			Pr.	X																					
			TI																						
		Formalizar proceso de seguridad de datos e información	Pe.	X																					
			Pr.	X																					

		Formalizar proceso de arquitectura	TI																		
			Pe.	X																	
			Pr.	X																	
		TI																			
		Pe.	X																		
		Pr.																			
	Cultura y transformación ⁵¹	Plan de comunicación	TI																		
			Pe.	X																	
			Pr.																		
		Plan de capacitación	Pe.	X																	
			Pr.																		
			TI																		
Gestión y control	Gestión y control	Indicadores de gestión y control	Pe.	X																	
			Pr.	X																	
			TI	X																	

5.2 Indicadores de seguimiento y control

Si bien durante la construcción de la propuesta de los programas y proyectos se destaca la formulación de indicadores de medición particulares, en este capítulo se establecerá el formato (tabla 30) y los contenidos mínimos necesarios para el seguimiento, tomando como soporte los planteamientos de Portilla (2016). Se estructura la metodología para realizar su construcción y poder aplicarla de forma transversal en cualquier momento.

Tabla 31 Formato definición de indicadores

Indicador		I Trim.	II Trim.	III Trim.	IV Trim.
Nombre	Definición del indicador.				
Fórmula	Expresión numérica o conceptual con la que se realizará la medición del indicador.				
Objetivo	Qué pretende medir el indicador definido, es decir, el impacto que se requiere controlar.				
Proceso	Las fases que el indicador está impactando.				
Umbral	Establecer un máximo, un mínimo o un parámetro aceptable				
Análisis	Revisión de los resultados y comportamientos de los indicadores en un determinado momento.				

⁵¹ Traducir las acciones planteadas en el capítulo 4.6.2 en un cronograma con responsables y tiempos.

En la siguiente tabla se ejemplifica un modelo de indicador asociado al gobierno de datos:

Tabla 32 Ejemplo definición y seguimiento de indicador

Indicador		I Trim.	II Trim.	III Trim.	IV Trim.
Nombre	Generación de ingresos antes y después de la puesta en marcha del modelo de GD.				
Fórmula ⁵²	Ingresos periodo actuales / ingresos período anterior.				
Objetivo	Medir el impacto del modelo de gobierno de datos en la generación de ingresos de la Corporación .				
Proceso	Contable - Financiero				
Umbral	15% de crecimiento después de la puesta en marcha del GD.				
Análisis	Los ingresos presentan o no crecimiento y se puede deducir si la implementación del modelo ha tenido o no impacto sobre los mismos.				

Los indicadores que se pueden aplicar en este tipo de ejercicio son:

Estratégicos:

- Crecimiento de ingresos derivados de la gestión de datos
- Cumplimiento de presupuesto derivado de la gestión de ingresos
- Nivel de adherencia del modelo de gobierno de datos (encuesta - anexo 2)
- Cantidad de iniciativas puestas en marcha / cantidad de iniciativas recibidas
- Gasto ejecutado gestión de datos / gasto total

Operativos:

- Cantidad de actividades reales / cantidad de actividades planeadas
- Cantidad de registros defectuosos / cantidad de registros correctos
- Cantidad de incidentes sin solucionar / cantidad de incidentes recibidos
- Cantidad de participantes en cada iniciativa / cantidad total de trabajadores
- Cantidad de charlas y capacitaciones generadas en un período determinado
- Volumen y cantidad de multas asociadas a la calidad de datos

⁵² Los indicadores pueden ser porcentajes o valores numéricos.

6 CONCLUSIONES

Si las empresas aspiran a tener continuidad en el tiempo, es indispensable re pensar la toma de decisiones, cuya base ahora son los datos. El desafío se encuentra en transformar los datos en respuestas, pero para lograr esto hay que administrarlos correctamente y garantizar la calidad, confiabilidad y disponibilidad de la información con que cuenta la corporación.

Cuando se habla de la nueva religión de los datos o el “datoismo” (Harari, 2016), se deduce que los temas asociados no serán una moda, sino más bien, que se convertirán en parte del quehacer empresarial, como una función adicional dentro de la operación. Así, la implementación un modelo de gobierno de datos, será un requerimiento para ser competitivo y no una simple iniciativa interna.

Se pueden concluir lo siguiente:

- En este documento se cimientan las bases como punto de partida para estructurar un modelo de gobierno para el sector de las Cajas de Compensación Familiar. Al realizar la mezcla de diferentes herramientas metodológicas se estructuró una propuesta general para el sistema.
- Por la diversidad de operaciones de estas 43 empresas, el marco de referencia unificado permitirá saber por dónde comenzar a abordar esta iniciativa; ya que las metodologías del mercado indican lo que se debe hacer, más no cómo hacerlo.
- En el modelo de madurez aplicado en la empresa seleccionada, se identifica que la estrategia de calidad de datos es la más relevante; y, por lo tanto, tiene un sentido de urgencia mayor a las demás iniciativas.
- El proceso de implementar un modelo de gobierno de datos no solucionará todos los problemas asociados a los datos, pero sí es el punto de partida para empezar a organizar los temas asociados a los mismos. De ahí surge la necesidad de lograr realizar priorizaciones y no abordar todas las iniciativas al mismo tiempo.
- Resulta crucial la definición de los elementos a incluir dentro del programa de gobierno de datos, es decir, qué es y qué no es gobierno; para no incluir componentes operativos que se deben solucionar a nivel de negocio, porque no tienen la connotación de estratégicos.
- La identificación de las audiencias claves dentro de la CCF es primordial para poner en marcha el programa, los dueños de los datos conocen la operación

y producción de los mismos y de esta manera se contribuye a asegurar el éxito.

- La decisión para dar cumplimiento a este requerimiento necesita una serie de pasos y planeación de actividades, en sí no funciona solo; las cuales se definieron así:
 - Análisis interno para el punto de partida, y la definición del estado futuro.
 - La definición de los factores de éxito, los componentes, los roles y responsabilidades y el modelo de operación, así como la cadena de valor y el ciclo de vida de los datos.
 - Establecer programas y proyectos con un cronograma claro, con indicadores de seguimiento y responsables encargados de llevar a buen término los objetivos establecidos.
 - Estrategias para impactar la cultura organizacional durante todo el proceso.

De esta manera, se abarcan todos los elementos necesarios para construcción de la hoja de ruta o plan de implementación, porque incluye todos los pilares que van a garantizar el éxito de la iniciativa.

Para cerrar la construcción de este marco de referencia se propone precisamente el mapa de ruta, incluyendo las etapas del proceso, la iniciativa, actividad, componente y el tiempo para desarrollar esta propuesta en 8 meses.

6.1 Trabajos futuros

Si bien el concepto de los datos no es nuevo, porque siempre ha estado presente en la evolución de la humanidad⁵³, el arte de formalizar las tareas de gobierno de datos, sí lo es. Se han ido estructurado metodologías que pretenden servir de faro para la ejecución en las instancias sociales, económicas y empresariales. Por lo tanto, no hay una receta única y cada ejercicio de gobierno se convierte en un entregable adaptado a las condiciones de los impulsores que llevan a iniciar este camino.

En este orden de ideas, el prototipo de gobierno de datos y las estrategias de gestión de datos, son apenas el primer paso, se trata de la formulación de la hoja de ruta para construir un verdadero modelo de una organización dirigida bajo los preceptos de los datos.

⁵³ Ejemplo planteado en Mayer-Schönberger (2013), donde hay registros contables de hace mas de 5.000 años.

Son muchos los elementos que estarían pendientes de incluir, en la medida que el programa vaya madurando y se permea en todos los niveles de la organización; nacen nuevas propuestas que requerirán el avance de otras capacidades y re pensar la estrategia actual. Se destacan los siguientes:

- Valoración de datos: se ha referenciado que los datos son el factor diferenciador para el éxito empresarial, pero es menester implementar una estrategia que, con una metodología clara, permita realizar la asignación de valor a los datos que se tienen y su impacto en la Corporación; lo cual dependerá de factores como el uso, importancia, relevancia, estrategia, etc., otorgado a los mismos.

Es importante destacar, tal como se mencionó en apartados anteriores, que el factor que mueve la economía ya no es el dinero, sino los datos, el capitalismo de los datos será el inductor para iniciar un proceso claro de valoración de datos.

- Medición de impactos: proceso de evaluar con una metodología clara y establecida, el impacto de una iniciativa de gestión de datos, es un trabajo relevante que demanda tiempo y recursos;

En sí, no hay iniciativa de gobierno que sea implementada si no tiene una injerencia directa en los objetivos estratégicos de la Corporación, cuya medición se debe realizar de manera puntual, una vez se haya formalizado el programa.

Se busca establecer formas de medir el impacto en los ingresos, en los nuevos negocios, en la mejora en la productividad e incluso en la cultura organizacional; y para poder monitorear estas incidencias, es necesario una medición formal con un alcance establecido.

- Las iniciativas de bodega de datos y operaciones, integración e interoperación, gestión de documentos y contenidos, inteligencia de negocios, Big Data y Ciencia de datos; se deben desarrollar y establecer como estrategias independientes; en este documento se trataron de forma implícita, pero deben ser visibles con inductores, responsables, cronograma y conceptos.
- Es conveniente abordar el diccionario de datos como un entregable independiente, entendido como el instrumento técnico que describe los datos; en este documento se plantea el glosario de negocio, pero no es suficiente para realizar un entendimiento completo de los datos desde los sistemas fuente hasta los repositorios de datos (productores y consumidores).

REFERENCIAS BIBLIOGRÁFICAS

- Ampuero, L., Alfaro, R., Raymundo, C., Domínguez, F. (2018). Modelo de madurez tecno-organizacional para la puesta en marcha exitosa de iniciativas de Data Governance. Perú. Escuela de Ingeniería de Sistemas, Universidad Peruana de Ciencias Aplicadas (UPC).
- Batini, C., M. Scannapieco. (2006) Data Quality: Concepts, Methodologies and Techniques. Data-Centric Systems and Applications. Berlin: Springer-Verlag Berlin Heidelberg.
- Bernal, L. (2018). Marco conceptual de la gobernanza de datos en el dominio de las redes inteligentes de Colombia. Recuperado de <https://cutt.ly/Ef5xyab>
- Calabrese, J., Esponda, S., Pasini, A., & Otros. (2019). Guía para evaluar calidad de datos basada en ISO/IEC 25012. Instituto de Investigación en Informática. Buenos Aires.
- Canals, Jordi. (2004). Crecimiento empresarial: personas y tecnología en la nueva economía. Revista Empresa y Humanismo, Vol. II, Nº 2/00, pp. 337-370
- Conpes 3920. (2018). Política nacional de explotación de datos (Big Data). Recuperado de <https://cutt.ly/6f5lm1Z>
- CMMI Institute (Capability Maturity Model Integration). (2014). Data Management Maturity Model (DMM). Version 1.1
- DAMA International. (2017). DMBOK (Data Management Body of Knowledge). 2nd Edition. Technics Publications LLC.
- DAMA internacional. (2010). La guía DAMA de los fundamentos para la gestión de datos. Publicaciones Técnicas.
- DAMA International. (2011). The DAMA dictionary of data management. 2nd Edition. Technics Publications LLC.
- Harari, Y.N. (2014). *Sapiens. De animales a dioses. Una breve historia de la humanidad*. Barcelona, España. Editorial Debate.
- Harari, Y.N. (2016). *Homo Deus. Breve historia del mañana*. Barcelona, España. Editorial Debate.
- Henderson J. C., Venkatraman N. (1999). Strategic alignment: Leveraging information technology for transforming organizations. IBM Systems Journal, Vol 32, No. 1
- ISO/IEC/ 25012. Disponible en <https://cutt.ly/QgnfcL8>
- Marchildon, P., Bourdeau, S., & otros (2018). Data governance maturity assessment tool: a design science approach. Boeck Supérieur
- Mayer-Schönberger V., Cukier K. (2013). Big Data. La revolución de los datos masivos. España. Turner Publicaciones S.L.
- Mayer-Schönberger V., Range T. (2019). La re-inención de la economía. El capitalismo en la era del Big Data. España. Turner Publicaciones S.L.
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2014). G.INF.06 Guía Técnica – Gobierno del dato. Recuperado de: <https://cutt.ly/Cf5xgXL>

- Morelos-Gómez, J., Fontalvo, T. (2014). Análisis de los factores determinantes de la cultura organizacional en el ambiente empresarial. *Revista Entramado*. Vol. 10 No. 1
- Muñoz, G., Quiroz, G. (2015). *Hacia un modelo de organización manejada por datos*. México.
- OCDE (Organización para la Cooperación y el Desarrollo Económicos). (2016) *Principios de gobierno corporativo de la OCDE y del G20*. Éditions OCDE, Paris. <http://dx.doi.org/10.1787/9789264259171-es>
- Osorio, M., Guerrero, C., González-Zabala, M. (2017). La gobernabilidad de los datos como apoyo en la gestión de datos de instituciones de educación superior. *Revista Espacios*. Vol. 38 (Nº 51) Año 2017. Pág. 11
- Portilla, J. (2016). Efectos de la integración de un programa de gobierno de datos con el sistema de gestión de calidad en Positiva Compañía de Seguros. Recuperado de <https://cutt.ly/hf5xPk0>
- Salim, I., Malone, M.S., Van-Geest, Y. (2016). *Organizaciones exponenciales*. España. Bubok Publishing S.L.
- Soares, S. (2010). *The IBM Data Governance Unified Process*. MC Press Online, LLC. USA.
- Soares, S. (2011). *Selling Information Governance to the Business*. MC Press Online, LLC. USA.
- Soares, S. (2014). *The Chief Data Officer Handbook for Data Governance*. MC Press Online, LLC. USA.
- Uribe, C. (2007). *Ciencia, Tecnología y Sociedad: Evolución y Revoluciones*. Medellín-Colombia. Instituto de Química. Universidad de Antioquia.
- Vallejos, E., Soto, M. (2019). Diseño del programa de Gobierno de Datos del Hospital General de Medellín. Recuperado de <https://cutt.ly/kf5xUT4>

ANEXOS

Anexo 1. Encuesta de análisis exploratorio y de medición de percepción.

Componentes	#	Pregunta	Opción				
			Totalmente de acuerdo	De acuerdo	Ni en acuerdo ni en desacuerdo	En desacuerdo	Totalmente en desacuerdo
Estrategia de gestión de datos	1	¿Existe apoyo visible y activo de la alta dirección para la estrategia de gestión de datos*?					
	2	¿La estrategia de gestión de datos está alineada con el direccionamiento estratégico?					
	3	¿Hay suficiente entendimiento y acuerdo entre las partes interesadas - directivos y operativos-, para apoyar un programa de gestión de datos sostenibles a largo plazo?					
	4	¿Los proyectos están alineados con la aplicación de una estrategia de gestión de datos?					
	5	¿Las capacidades del personal y recursos físicos están diseñadas para crear y dirigir un programa de gestión de datos?					
	6	¿Hay compromiso de proporcionar el entrenamiento que permita la madurez de un programa de gestión de datos?					
Comunicación	7	¿Existen condiciones, normas y procedimientos para la comunicación de un programa para la gestión de datos?					
	8	¿Se mantiene a la organización informada acerca del programa de gestión de datos y proyectos de gestión asociados?					
	9	¿Se logra la comunicación bidireccional entre los negocios, tecnología, programa de gestión de datos y alta dirección acerca de las prioridades en la gestión de datos, enfoques y entregas?					
Función de la gestión de datos	10	¿La función de gestión de datos está definida de tal manera que es clara a todas las partes interesadas?					
	11	¿La función de gestión de datos está alineada con la estrategia de gestión de datos, demostrada por medidas y métricas?					
	12	¿La alta dirección participa activamente en el diseño y supervisión de la función de gestión de datos?					
Gestión del gobierno	13	¿El gobierno de datos proporciona mecanismos para facilitar la colaboración y toma de decisiones a través de líneas de negocio y funciones de TI?					

	14	¿La estructura de gobierno de datos delimita claramente las responsabilidades de los dominios de datos*?					
	15	¿La organización define roles y responsabilidades y asegura que todas las partes interesadas participen?					
	16	¿Tiene el gobierno de datos un mecanismo para la definición y resolución de prioridades?					
	17	¿El gobierno de datos proporciona efectivamente un proceso para definir, escalar y resolver problemas?					
	18	¿Tiene el gobierno de datos un patrocinador visible que apoye la iniciativa?					
	19	¿La organización ha establecido un programa de cumplimiento efectivo en todo el ciclo de vida de los datos?					
	20	¿Tiene la organización un proceso para examinar la estructura y las actividades de gobierno?					
	21	¿Hay una formación adecuada para el personal involucrado en la gestión de datos?					
	22	¿Existe un proceso para llevar a cabo el cumplimiento de las decisiones de los componentes del gobierno de datos?					
Glosario o diccionario de datos de negocio	23	¿Existe una política que obliga el uso y referencia al glosario de datos de negocio*?					
	24	¿Está el glosario de datos de negocios promulgado y accesible para todos los interesados?					
	25	¿Es el glosario de datos el primer paso en el diseño de aplicaciones y proyectos de datos?					
	26	¿La organización realiza referencias cruzadas y mapeo de términos específicos (sinónimos, glosarios unidad de negocio, atributos lógicos, elementos físicos de datos, etc.) para lograr términos de negocio estandarizados?					
	27	¿Se mejora el glosario de negocio de la organización y se gestionan los cambios y adiciones?					
	28	¿El gobierno de datos participa en la creación, aprobación, gestión, y la actualización de los términos de negocio?					
	29	¿Existe un proceso de cumplimiento implementado para asegurar que las unidades de negocio y proyectos están aplicando correctamente términos de negocio?					
	30	¿La organización emplea un proceso definido para proporcionar retroalimentación sobre los términos de negocio?					

	31	¿Existe glosario de datos actualizado para todas los sistemas transaccionales?					
	32	¿Es accesible el glosario de datos a todas las partes interesadas a nivel técnico / tecnológico?					
	33	¿Existe modelado de datos en los sistemas transaccionales que permitan tener documentación, flujos y diagramas de arquitectura?					
Gestión de metadatos	34	¿Existe estrategia definida de metadatos*, alineada con el programa de gobierno de datos y estándares externos?					
	35	¿Existe análisis e identificación de metadatos para su inclusión dentro de la repositorio de metadatos definido?					
	36	¿Todas las partes interesadas participan en la definición de las categorías de metadatos y propiedades?					
	37	¿Existe método para el desarrollo y la evaluación de las normas y procesos para la gestión de metadatos?					
	38	¿Existe método para mantener o actualizar el repositorio de metadatos?					
	39	¿Los roles y responsabilidades están claramente definidos para la captura, actualización y uso de metadatos?					
Estrategia de la calidad de datos	40	¿La calidad de datos se destaca en todas las iniciativas de gestión de datos?					
	41	¿Se mide la evolución del programa de calidad de datos?					
	42	¿Existe una unidad de organización que es responsable de mantener la calidad de los datos?					
	43	¿Se generan iniciativas de calidad de datos en las unidades de la organización?					
	44	¿Hay roles, responsabilidades y rendición de cuentas claramente definidos para fomentar mejora de la calidad de los datos?					
	45	¿Se promulga y comunica la estrategia de calidad de datos?					
	46	¿La estrategia de calidad de los datos describe claramente los objetivos, las políticas y procesos?					
	47	¿Está la calidad de datos integrada con el ciclo de vida del desarrollo de tecnología?					
	48	¿Se realizan esfuerzos de mejora en la calidad de datos integrada con los procesos de negocio?					
De perfiles de datos	49	¿Tiene la organización un método estándar para realizar perfiles de datos*?					
	50	¿La organización capacita y tiene personal con experiencia en los datos herramientas y técnicas de perfilado?					

	51	¿En la organización se aplican modelos estadísticos para analizar perfiles de datos?					
	52	¿El perfilado de datos está basado en eventos definidos y programados?					
Evaluación de la calidad de datos	53	¿Las técnicas de evaluación de calidad de datos son estándar y se cuentan con métodos documentados?					
	54	¿Se llevan a cabo evaluaciones de calidad de datos?					
	55	¿Las reglas de calidad de datos son estándar y se desarrollan para los atributos de datos básicos?					
	56	¿Se emplean motores de reglas de calidad o herramientas de evaluación?					
	57	¿Existen impactos de costos de los problemas de calidad de datos y se utilizan como entrada a las prioridades de mejora de calidad de datos?					
Limpieza o depuración de datos	58	¿Tiene la organización un conjunto de procesos de limpieza de datos (automático y manual) para resolver los problemas de calidad de datos?					
	59	¿Existe un proceso definido para verificar la corrección y la evaluación de la efectividad de la depuración de datos?					
	60	¿La organización limpia los registros duplicados?					
	61	¿Se implementan correcciones y puntos de control en la fuente de captura de datos?					
	62	¿Los datos siguen los procesos de limpieza a través del análisis de las causas fundamentales?					
	63	¿Se definen umbrales de calidad y los límites de tolerancia?					
	64	¿La organización ha desplegado un conjunto de herramientas consistentes para soportar datos limpieza?					
	65	¿El retorno de la inversión incorpora los costos de limpieza de datos?					
	66	¿En la organización se aplican consideraciones operativas y de reputación para determinar qué actividades de limpieza de datos se deben financiar?					
67	¿La organización define y supervisa los procesos de limpieza de datos?						
Operaciones de datos	68	¿Existen procesos de solicitud, captura, evaluación y verificación de datos con las partes interesadas?					
	69	¿Existen requisitos de datos asignados a los objetivos de negocio?					
	70	¿Existen definiciones de datos estándar, así como representaciones lógicas y físicas?					
	71	¿Se definen actividades, hitos y productos para el mapeo de procesos de los datos creados?					

	72	¿La organización ha establecido funciones y responsabilidades claras para la creación y mantenimiento de un modelo de datos de negocio*?					
	73	¿Los métodos de modelado de procesos son estándar, y se cuentan con herramientas para modelar y definir los procesos de negocio?					
	74	¿Se capturan necesidades de abastecimiento de datos, validados y priorizados?					
	75	¿Existe un mecanismo que asegura las necesidades de abastecimiento?					
	76	¿Los atributos de datos se asignan a las fuentes de datos y a las aplicaciones posteriores?					
	77	¿Se gestiona el proceso de selección de la fuente de datos?					
	78	¿Existe un proceso repetible para la gestión de la operación de datos?					
Plataforma y arquitectura	79	¿Se genera enfoque de la arquitectura de organización hacia los activos de información?					
	80	¿Hay seguimiento constante al enfoque arquitectónico de los activos de información?					
	81	¿Se emplea el método de sincronización para la racionalización, consolidación y eliminación de datos duplicados?					
	82	¿Tiene la organización un conjunto de tecnología de datos aprobado, para aplicar modificaciones, adiciones, y actualizaciones?					
	83	¿Se definen las categorías de los estándares requeridos para la meta de la organización en la arquitectura de datos?					
	84	¿Determina la organización la necesidad de la estrategia de tecnología para el desarrollo de acceso a datos estándar y de aprovisionamiento?					
	85	¿Se aprueban los modelos de datos, se mantienen y son gobernados?					
	86	¿La organización ha definido arquitectónicamente, el acceso de datos estándar y criterios para determinar qué métodos se aplican?					
	87	¿Las fuentes de datos tienen niveles de autoridad definidos, seleccionados, e integrados?					
	88	¿Existe un modelo de seguridad de la información definido?					
Integración de datos	89	¿Tiene la organización un proceso para hacer adquisición y abastecimiento de datos externos?					
	90	¿Se evalúan las necesidades de consolidación de datos?					
	91	¿Se minimiza la redundancia de datos?					
	92	¿Se consolidan los datos de manera efectiva?					

	93	¿Existen normas de integración de datos y son revisados, supervisados, aprobados y aplicados?					
	94	¿Se describen los procesos de cumplimiento para la integración de datos?					
	95	¿Se definen umbrales de calidad de datos aplicados a las fuentes de captura?					
	96	¿Existen procesos para identificar los datos faltantes y se hace el seguimiento contra defectos o faltantes?					
	97	¿Se posee el personal adecuado para controlar, gestionar y sostener calidad de los datos?					
Datos históricos	98	¿Existen estándares de arquitectura y convenciones aplicadas a la estructura y la gestión de los datos históricos, con reglas definidas y gobernadas?					
	99	¿Existe proceso para la retención de datos por el tiempo definido?					
	100	¿Se mantiene la integridad de los datos archivados?					
	101	¿Existe un enfoque consistente para la recuperación e integración de archivado de los datos históricos con los datos actuales?					
	102	¿Se realiza auditoría de cambios en los datos, los cuales se monitorean y administran?					
	103	¿Existen consideraciones y políticas para determinar cuándo los datos archivados pueden ser eliminados?					
Procesos de soporte	104	¿Existen medidas y análisis para determinar si las metas de gestión de datos se están cumpliendo?					
	105	¿Se definen procesos de medir, analizar y controlar los datos?					
	106	¿Las medición y análisis están integrados en la gestión del proceso de datos?					
	107	¿Se mantienen procesos, métodos, procedimientos, políticas y estándares en torno a los datos?					
	108	¿Se mide el rendimiento del proceso de gobierno de datos?					
	109	¿Se analizan los problemas de calidad de datos de tendencia negativa?					
	110	¿Todas las partes interesadas tienen visibilidad en la calidad del proceso de gobierno de datos?					
	111	¿La organización conoce la cantidad de riesgo asociada al manejo de datos?					
	112	¿Ha identificado la organización un mapa de riesgos asociados a los datos?					
	113	¿Tiene la organización supervisión periódica de los riesgos y realiza actualizaciones?					
	114	¿Se gestionan los datos y se controlan a través del ciclo de vida de los datos?					

Anexo 2. Encuesta percepción para identificar adherencia del programa de gobierno de datos.

Estrategia	Objetivo	Pregunta	Opciones
Gobierno de datos	Medir el nivel de adherencia del programa.	¿Conoce el programa de gobierno de datos?	1- Totalmente de acuerdo 2- De acuerdo 3- Ni en acuerdo ni en desacuerdo 4- En desacuerdo 5- Totalmente en desacuerdo
		¿Aplica los enunciados del decálogo de gobierno?	
		¿Se usan los datos para todas las decisiones estratégicas, tácticas y operativas?	
		¿Están las capacidades de la organización orientadas a lograr la evolución del GD?	
		¿Los roles y las responsabilidades de las partes interesadas son suficientemente claros?	
		¿Las iniciativas alrededor de los datos se canalizan fácilmente al comité de GD?	
		¿Considera que la estructura organizacional obedece a una empresa manejada por datos?	
		¿El diccionario de negocio es usado en todas las iniciativas de datos de la empresa?	
		¿Conoce las actividades del semillero de gestión de datos de la empresa?	
		¿Le son familiares los términos asociados al gobierno (roles, responsabilidades, metadatos, datos de referencia)?	
		¿Se identifican y gestionan los riesgos asociados a los datos del área?	

Anexo 3. Decálogo del gobierno de datos.

Ítem	Concepto
1	El gobierno de datos hace parte integral de la estrategia de gestión, incluida dentro de los planes estratégicos de la Corporación.
2	El comité de gobierno de datos es la máxima autoridad decisiva en la organización.
3	Las decisiones estratégicas, tácticas y operativas se toman con base en los datos.
4	Los datos se encuentran al servicio de todos y para todos, garantizando la democratización de los mismos.
5	Las políticas y procesos definidos para la gestión de datos se aplicarán sin excepciones.
6	El gobierno de datos es una iniciativa centrada en el negocio, tecnología y otras áreas son el soporte y apoyo.

7	El desarrollo y éxito del programa dependerá del esfuerzo y la colaboración de todas las áreas involucradas.
8	Los datos deben cumplir la función principal de generar valor y así se deben realizar los esfuerzos para lograrlo.
9	Las iniciativas de gestión de datos provienen de cualquier nivel de la organización.
10	Se usarán y desarrollarán las herramientas tecnológicas determinadas para la explotación y exploración de datos.