

El futuro digital
es de todos

Gobierno
de Colombia
MinTIC

G.INF.03 Guía técnica de Información-Ciclo de vida del dato

Ministerio de Tecnologías de la Información y las Comunicaciones

Viceministerio de Economía Digital

Dirección de Gobierno Digital

Subdirección de Estándares y Arquitectura de TI

Equipo de trabajo

Sylvia Cristina Constaín Rengifo – Ministra de Tecnologías de la Información y las Comunicaciones

María Isabel Agudelo - Viceministra de Economía Digital (e)

Claudia Patricia Pico Quintero – Directora de Gobierno Digital

Leydi Viviana Cristancho Cruz – Subdirectora de Estándares y Arquitectura TI

Martin Antonio Orjuela Velasco – Equipo de la Subdirección de Estándares y Arquitectura de TI

Nicolás Sánchez Barrera – Equipo de la Subdirección de Estándares y Arquitectura de TI

Gamaliel Andrés Silva Ortiz – Equipo de la Subdirección de Estándares y Arquitectura de TI

Anyelina Lalage Cáceres Reyes – Equipo de la Subdirección de Estándares y Arquitectura de TI

Daniel Castillo Bernal – Equipo de la Subdirección de Estándares y Arquitectura de TI

Versión	Observaciones
Versión 1.0 Diciembre 30 de 2014	Emisión
Versión 1.1 Octubre de 2019	Actualización Gobierno Digital

Comentarios, sugerencias o correcciones pueden ser enviadas al correo electrónico: gobiernodigital@mintic.gov.co

Construcción del PETI – Planeación para la Transformación Digital



Este documento de la Dirección de Gobierno Digital se encuentra bajo una [Licencia Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).

Versión 1.0

María Isabel Mejía Jaramillo

Viceministra de Tecnologías y Sistemas de la Información
Ministerio de Tecnologías de la Información y las Comunicaciones

Jorge Fernando Bejarano Lobo

Director de la Dirección de Estándares y Arquitectura de TI
Ministerio de Tecnologías de la Información y las Comunicaciones

Asesores del Ministerio de Tecnologías de la Información y las Comunicaciones

Claudia Milena Rodríguez Álvarez
Carlos Arturo Merchán Herrera
Esteban Armando Gaviria García

Asesores de la Corporación Colombia Digital

Javier Orlando Torres Páez
Deicy Alexandra Parra Chaux
Diego Antonio Campos Cáceres
Jaime Leonardo Bernal Pulido
Leydi Viviana Cristancho Cruz

Medios Digitales

María Astrid Toscano Villán
Ricardo Rojas Ortíz
Jhon Henry Munevar Jiménez

UT Everis Tecnom

Alberto Pizarro Carrasco
Gerardo Antonio Moreno
Martha Lucía Parra
Martha Patricia Naranjo Becerra
David Fernando de la Peña Santana
Lucio Augusto Molina Focazzio
Silvia María Fernández Coello

Karin Xiomara Marroquín
Maribel Ariza Rojas
Ramiro Andrés Delvasto
Diego Ordóñez
Edgar Esquiaqui
Ricardo Abad Chacón Ibama
Juliana Botero Iragorri
Juan Pablo Sequera España

TABLA DE CONTENIDO

PÁG.

TABLA DE CONTENIDO	5
LISTA DE IMAGENES	7
LISTA DE TABLAS	8
ABREVIATURAS Y ACRÓNIMOS	9
1 INTRODUCCIÓN	12
2 OBJETIVOS	13
3 ALCANCE DE LA GUÍA.....	14
4 LINEAMIENTOS DEL MARCO DE REFERENCIA DE AE ASOCIADOS	15
5 ¿QUÉ ES EL CICLO DE VIDA DEL DATO?	16
6 ¿QUÉ FLUJO DE PASOS DEBE DESARROLLAR UNA INSTITUCIÓN PARA GESTIONAR EL CICLO DE VIDA DEL DATO?.....	21
7 ¿QUÉ DETALLE DE PASOS DEBE DESARROLLAR UNA INSTITUCIÓN PARA GESTIONAR EL CICLO DE VIDA DEL DATO?.....	23
7.1 Paso 1: Generar	23
7.2 Paso 2: Evaluar.....	24
7.3 Paso 3: Clasificar.	26
7.4 Paso 4: Analizar.	28
7.5 Paso 5: Rediseñar los procesos	30
7.6 Paso 6: Consolidar	33
7.7 Paso 7: Conservar	34



8	GLOSARIO.....	37
9	REFERENCIAS.....	38

LISTA DE IMAGENES

	PÁG.
Ilustración 1. Ejes para el ciclo de vida del dato.....	16
Ilustración 2. Factores claves de éxito para una adecuada gestión del ciclo de vida del dato	18
Ilustración 3. Pasos para desarrollar e implementar la gestión del ciclo de vida del dato	21

LISTA DE TABLAS

	PÁG.
Tabla 1. Paso 1: Generar.....	24
Tabla 2. Paso 2: Evaluar.....	26
Tabla 3. Paso 3: Clasificar.....	28
Tabla 4. Paso 4: Integrar MDM	30
Tabla 5. Paso 5: Manipular	32
Tabla 6. Paso 6: Manipular	34
Tabla 7. Paso 7: Conservar	36

ABREVIATURAS Y ACRÓNIMOS

Abreviatura / acrónimo	Descripción
AQR	Reporte de Calidad para Arquitectura. (Architecture Quality Report).
BPM	Gestión de Procesos de Negocio (Business Process Management).
CCC	Campos Críticos de Calidad.
CMM	Modelo de Madurez de Capacidades (Capacity Maturity Model).
CRUD	Crear, Leer, Actualizar y Eliminar. CRUD (Create, Read, Update, Delete). Se usa para referirse a las funciones básicas en bases de datos o la capa de persistencia en un software.
DBMS	Sistema Manejador de Base de Datos (Data Base Management System).
DBQR	Reporte de Calidad de Base de Datos (Data Base Quality Report).
DGQR	Reporte de Calidad para Gobernanza de Datos (Data Governance Quality Report).
DQR	Reporte de Calidad para Datos (Data Quality Report).
DRDA	Estándar de Arquitectura Distribuida para Bases de Datos Relacionales del Open Group (Data Relational Database Architecture).
DUQR	Reporte de Calidad para Uso de Datos (Data Use Quality Report).
ECM	Gestión de Contenido Empresarial (Enterprise Content Management).
ERP	Planeamiento de Recursos de la Institución (Entreprise

	Resource Planning)
ETL	Extracción, Transformación y Carga (Extract, Transform and Load) en sistemas de información y almacenes de datos.
ICT	Infraestructura Común de Telecomunicaciones.
ILM	Gestión del Ciclo de Vida de Información (Information Lifecycle Management).
IQR	Reporte de Mejora de Calidad (Improvement Quality Report).
KPI	Indicador Clave de Proceso (Key Performance Indicator).
LOPD	Ley Orgánica de Protección de Datos.
MDA	Arquitectura Orientada al Modelo (Model-Driven Architecture).
MDM	Gestión de Datos Maestros. (Master Data Management)
MinTIC	Ministerio de Tecnologías de la Información y las Comunicaciones.
MRP	Planeación de Recursos de Manufactura. (Materials Requirement Planning)
MOE	Ambiente Operativo Gestionado. (Managed Operating Environment)
PI	Plan de Implementación.
PQR	Reporte de Calidad para Procesos. (Process Quality Report)
PAAC	Proceso para la Adopción y Aceleración del Cambio.
QR	Mapa de Ruta de Calidad. (Quality Roadmap)
ROI	Retorno de la Inversión. (Return of Investment)
SLAs	Acuerdos de Nivel de Servicio. (Service level Agreement)
SOA	Arquitectura Orientada a Servicios. (Service Oriented Architecture)
SOE	Ambiente Operativo Estandarizado. (Standard Operating Environment)
TI	Tecnología de Información.

TIC	Tecnologías de la Información y la Comunicación.
-----	--

1 INTRODUCCIÓN

A continuación, se presentan los objetivos y el alcance de la guía, así mismo se identifican los lineamientos del Marco de Referencia de AE para la Gestión de TI, que apoyan la misma.



2 OBJETIVOS

Los siguientes son los principales objetivos:

- Proveer a las instituciones del sector público con una guía para la gestión del ciclo de vida de los datos.
- Facilitar la implementación de los lineamientos del Marco de Referencia de AE para la Gestión de TI, relacionados con el ciclo de vida del dato.
- Facilitar la planeación del ciclo de vida para diferentes componentes de información de características específicas como, por ejemplo, el ciclo de vida de la información geográfica, multimedia, etc.

3 ALCANCE DE LA GUÍA

La presente guía tiene el siguiente alcance:

- Definir el conjunto de pasos y/o actividades que permitan gestionar el ciclo de vida del dato en las instituciones.
- Entender el nivel de madurez del dato, considerando la especificación y el aprovisionamiento de un dato administrado hacia el autoservicio.

4 LINEAMIENTOS DEL MARCO DE REFERENCIA DE AE ASOCIADOS

Los siguientes lineamientos del dominio de Información, del Marco de Referencia de AE para la Gestión de TI, son apoyados de manera directa por la guía:

- LI.INF.06 Lenguaje Común de Intercambio de componentes de información.
- LI.INF.08 Publicación de los servicios de intercambio de componentes de información.
- LI.INF.05 Definición y caracterización de la información georreferenciada.
- LI.INF.06 Lenguaje común de intercambio de Componentes de Información.

5 ¿QUÉ ES EL CICLO DE VIDA DEL DATO?

La gestión del ciclo de vida del dato se enfoca en asegurar el cubrimiento del enmascaramiento, reducción, archivado, y generación (para prueba) de los datos en la institución, soportando el nivel de autoservicio comprometido. Esta gestión es una práctica basada en políticas que se encarga del flujo de los datos de los sistemas de información a través de su ciclo de vida: desde la creación y el almacenamiento inicial, hasta el momento cuando se convierte en obsoleto y es eliminado. Adicionalmente, el ciclo de vida del dato es uno de los ámbitos principales del gobierno del dato (ver guía técnica) en la institución del sector público.

En el Marco de Referencia de AE, se resaltan los elementos del ciclo de vida como ejes, los cuales se describen a continuación:

EJE	DESCRIPCION	OBJETIVO
Archivado de datos históricos	Archivado coherente de información histórica en soportes IT más económicos, accesibles igualmente desde los aplicativos de producción.	Reducción del costo operativo de producción.
Enmascaramiento de entornos	Enmascaramiento coherente entre los datos sensibles, generados y procesados por los diferentes sistemas de la entidad, teniendo en cuenta las políticas de seguridad, privacidad y datos abiertos; establecidos por el Estado colombiano.	Cumplimiento normativo y protección del activo del dato.
Reducción de entornos	Reducción coherente del volumen de los entornos no productivos, para disminuir el coste IT asociado.	Reducción del coste asociado a los entornos no productivos.
Generación de datos de prueba	Descarga y enmascaramiento de datos coherentes de Producción en entornos de desarrollo con objeto de disponer de casos de prueba, soporte a la resolución de incidencias o externalización de los entornos de desarrollo.	Disminución de costos de desarrollo.

Ilustración 1. Ejes para el ciclo de vida del dato

Fuente. Elaboración propia.

Un enfoque de reducción es la identificación de las estrategias para existencia de entornos, a partir de los conceptos de la actividad misional, de las necesidades funcionales de pruebas y por último, de las dependencias técnicas de los modelos de datos.

Las características de los procesos de reducción son:

- Reproducible.
- Asegura coherencia (entre tablas que compartan información sensible).
Permite utilizar un porcentaje de reducción configurable.
- Permite múltiples criterios de reducción definidos en función de sentido de la actividad misional.

Considerando un nivel de madurez superior para el ciclo del dato, se define un dato orientado al autoservicio como aquel que: es predictivo, está alineado con procesos de negocio, viene dirigido a una audiencia por su contenido, responde vía servicio por demanda, y usa infraestructura requerida por demanda. El proceso para gestión del ciclo de vida del dato requiere llevar el dato hacia el autoservicio. También, el enfoque de autoservicio en el ciclo de vida del dato requiere del desarrollo y gestión del enmascaramiento, la reducción, y el archivado de los datos de la institución.

En la siguiente imagen, se pueden apreciar los factores claves para hacer una adecuada gestión del ciclo de vida del dato; muchos de estos factores están relacionados con la coordinación entre las diferentes áreas, la gestión de ambientes, y la metodología de trabajo, entre otros aspectos.



Ilustración 2. Factores claves de éxito para una adecuada gestión del ciclo de vida del dato

Fuente. Elaboración propia

A continuación, se describen los factores de éxito anteriormente presentados:

- **Modelo de relación y gestión de la demanda:** El establecimiento establecer modelos de gestión de la demanda de los datos, que optimice esfuerzos, y esté orientado a las necesidades reales y los objetivos estratégicos de las instituciones.
- **Definición procesos y servicios:** establecer formalmente los procesos y servicios que generan y modifican datos, y que dan soporte a las diferentes áreas de la institución, prestando especial atención a las relaciones entre

dichas áreas, e incluyendo los procesos y servicios propios de la Dirección de Tecnologías y Sistemas de la Información o quien haga sus veces.

- Reutilización de procesos: estandarizar los procesos que generan y modifican los datos con el fin de formalizarlos y hacer posible su reutilización.
- Parametrización y trazabilidad de los procesos: contar con procesos de Extracción, Transformación y Carga (ETL), parametrizables y que incorporen funcionalidades que ayuden a realizar una adecuada trazabilidad de las operaciones implementadas.
- Procesos re-ejecutables en el tiempo: incluir en los procesos, características que permitan la realización de ejecuciones puntuales, a demanda o planificadas.
- Coherencia con los sistemas centrales: asegurar la alineación de los procesos definidos para el ciclo de vida, con la forma en que gestionan los datos, los sistemas de información de apoyo de la institución (Gestión de la Relación con el Cliente -Customer Relationship Management – CRM, el Planeamiento de Recursos de la Institución - Enterprise Resource Planning – ERP, entre otros).
- Gestión de ambientes: contar con los diferentes ambientes que permitan realizar las pruebas correspondientes, y aseguren el adecuado manejo de los datos, respetando las políticas y lineamientos de la institución, respecto a la seguridad y privacidad de estos datos.
- Evolución de los modelos: definir estrategias de evolución para los modelos, que den soporte a la misión y a los objetivos estratégicos de la institución, acorde a las condiciones cambiantes del entorno.
- Enmascaramiento homogéneo: definir un método homogéneo para el enmascaramiento de los datos, tanto para el proceso como para las máscaras de entrada.

- Aleatoriedad por campo: establecer procesos de gestión de la aleatoriedad de los campos, teniendo en cuenta buenas prácticas (aleatorio, determinista, difuminación, sustitución, a medida).
- Recopilación de evidencias: contar con mecanismos que permitan registrar y recopilar evidencias del proceso de enmascaramiento de los datos, dichas evidencias serán objeto de análisis en las auditorías y el control interno de las instituciones.
- Custodia de RAW DATA: establecer mecanismos para la protección de los datos primarios (RAW DATA), que aseguren su integridad, confidencialidad y disponibilidad, acorde a las políticas de seguridad y privacidad establecidas por la institución.
- Coordinación en las ejecuciones: definir procedimientos que soporten la coordinación de las diferentes ejecuciones y pruebas con los responsables tanto de los datos, como de los sistemas de información que los generan, modifican o usan.
- Procedimientos de marcha atrás: establecer procedimientos formales de marcha atrás (Tipo rollback), que como mínimo, sean tenidos en cuenta en las operaciones que manejan los datos, que la institución ha establecido como críticos.
- Identificación de campos no triviales: identificar los campos de datos que contienen información compleja, susceptible de ser tratada, como pueden ser imágenes, documentos, audio, entre otros.

6 ¿QUÉ FLUJO DE PASOS DEBE DESARROLLAR UNA INSTITUCIÓN PARA GESTIONAR EL CICLO DE VIDA DEL DATO?

A fin de desarrollar e implementar la gestión del ciclo de vida del dato en la institución, se considera como mínimo ejecutar los siguientes pasos:

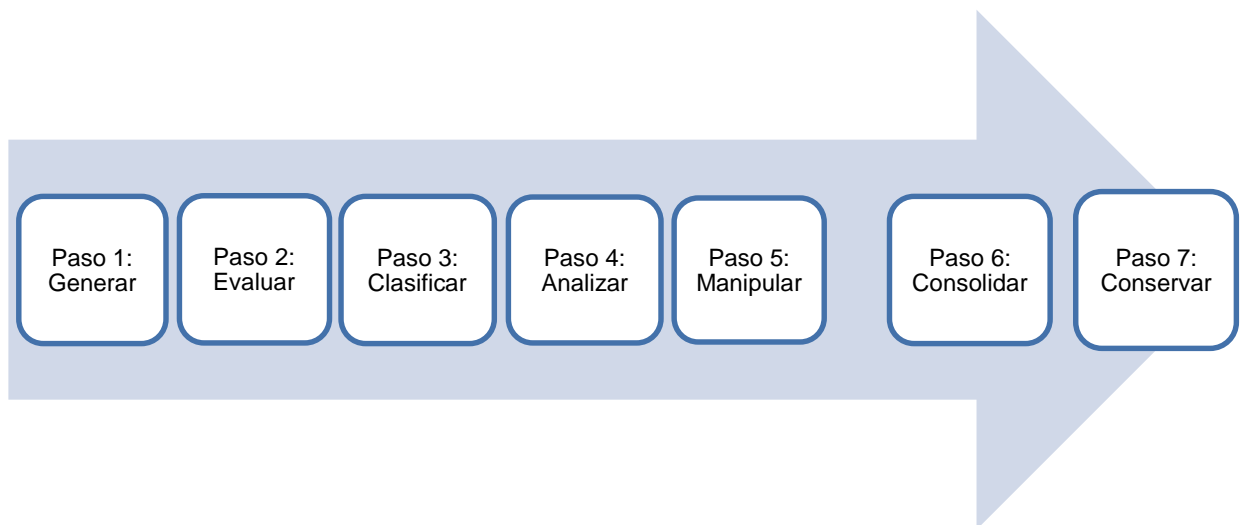


Ilustración 3. Pasos para desarrollar e implementar la gestión del ciclo de vida del dato

Fuente. Elaboración propia

En la imagen anterior, se ve cómo el flujo alrededor del ciclo de vida del dato inicia con la generación de un nuevo dato a gestionar, avanza con el estudio de las necesidades para su creación, continua con la categorización de características en su ciclo de vida, progresa con el análisis de requerimientos para mantener entornos del dato, avanza con la modificación realizada en sistemas y plataformas para

soporte de este ciclo, continua con el agrupamiento de entornos no productivos del dato, y finaliza, con la conservación en medio persistente de todos los atributos y reducciones asociadas al dato.

7 ¿QUÉ DETALLE DE PASOS DEBE DESARROLLAR UNA INSTITUCIÓN PARA GESTIONAR EL CICLO DE VIDA DEL DATO?

7.1 Paso 1: Generar

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 1.

PASO1: Generar	
OBJETIVOS	<ul style="list-style-type: none"> Identificar el origen del dato cuando entra en gestión del ciclo de vida. Definir la especificación del dato para soportar el ciclo de vida.
ENTRADAS	SALIDAS
<ul style="list-style-type: none"> Sistema de información origen del dato. Información recolectada desde el origen del dato (interno/externo). Especificación de nueva metadata para tipo de dato 	<ul style="list-style-type: none"> Descripción detallada del dato detectado (negocio) Metadata describiendo el origen del dato (arquitectura)
ACTIVIDADES	

<ul style="list-style-type: none"> ■ Arquitectura del dato <ul style="list-style-type: none"> ○ Identificar contexto de origen del dato, incluyendo sistemas, áreas organizacionales, normativa, requerimiento de negocio, etc. ○ Generar metadata del dato con la especificación entregada dentro de la mejor práctica de arquitectura. ○ Definir información para gestión del ciclo de vida del dato utilizando un lenguaje común de intercambio. ■ Autoservicio del dato <ul style="list-style-type: none"> ○ Identificar las operaciones que vuelven predictivo al dato. ○ Identificar los puntos e donde se alinea el dato con procesos de negocio. ○ Identificar las audiencias involucradas de acuerdo al contenido del dato. ○ Identificar las necesidades de servicio por demanda para el dato. ■ Identificar los requerimientos de infraestructura por demanda para conservar el dato.
--

Tabla 1. Paso 1: Generar

7.2 Paso 2: Evaluar

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 2.

PASO2: Evaluar	
OBJETIVOS	<ul style="list-style-type: none"> ■ Evaluar la pertinencia y calidad inicial del dato para gestionar ciclo de vida. ■ Tomar la decisión de almacenar o descartar el dato en ciclo de vida..

ENTRADAS	SALIDAS
<ul style="list-style-type: none"> ▪ Metadata del dato detectado (arquitectura) ▪ Descripción detallada del dato detectado (negocio) 	<ul style="list-style-type: none"> ▪ Aprobación de dato para almacenamiento o descarte. ▪ Reporte de características de calidad encontradas en dato evaluado para ciclo de vida. ▪ Acciones para especificar adicionalmente el dato antes de iniciar gestión del ciclo de vida.
ACTIVIDADES	
<ul style="list-style-type: none"> ▪ Calidad del dato <ul style="list-style-type: none"> ○ Evaluar criterios y campos críticos para calidad del dato. ○ Evaluar situaciones que implican depuración del dato. ○ Evaluar capacidad para gestión de calidad del dato. ▪ Enmascaramiento de datos <ul style="list-style-type: none"> ○ Evaluar situaciones que implican enmascaramiento del dato. ▪ Reducción (entornos o datos) <ul style="list-style-type: none"> ○ Evaluar situaciones que implican reducción del dato y sus entornos. ▪ Archivado de datos <ul style="list-style-type: none"> ○ Evaluar situaciones que implican conservación del dato en sus varios entornos. ▪ Autoservicio del dato <ul style="list-style-type: none"> ○ Evaluar como soportar las operaciones que vuelven predictivo al dato. ○ Evaluar los criterios de oportunidad, valor, y uso en donde se alinea el dato con procesos de negocio. ○ Evaluar la interacción esperada con las audiencias involucradas para el contenido del dato. ○ Evaluar el modelo de servicio por demanda para el dato. 	

<ul style="list-style-type: none"> ○ Evaluar el cumplimiento de los requerimientos de infraestructura por demanda para conservar el dato (existe capacidad, avance tecnológico, integración de plataformas, etc.) ■ Reporte del dato <ul style="list-style-type: none"> ○ Generar reporte del dato evaluado para ciclo de vida. ○ Aprobar el almacenamiento del dato para inicio de la gestión del ciclo de vida. En caso contrario, soportar la decisión para su descarte. ■ Acciones del dato <ul style="list-style-type: none"> ○ Identificar nuevas acciones para mejorar la evaluación sobre el ciclo de vida del dato. ○ Generar reporte de las nuevas acciones aprobadas. ○ Iniciar la implementación de estas acciones para mejorar el ciclo de vida del dato.
--

Tabla 2. Paso 2: Evaluar

7.3 Paso 3: Clasificar.

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 3.

PASO3: Clasificar	
OBJETIVOS	<ul style="list-style-type: none"> ■ Clasificar para ciclo de vida según la naturaleza y metadata del dato (origen). ■ Aplicar requerimientos del ciclo de vida orientados al autoservicio del dato.
ENTRADAS	SALIDAS

<ul style="list-style-type: none"> ▪ Descripción detallada del dato detectado (negocio). ▪ Metadata del dato detectado (arquitectura). ▪ Clasificación interna definida para los datos. ▪ Reporte de características de calidad encontradas en dato evaluado para ciclo de vida. ▪ Resultado de acciones adicionales para especificar dato e incorporarlo en ciclo de vida. 	<ul style="list-style-type: none"> ▪ Dato clasificado por metadata y con orientación al autoservicio. ▪ Referencia única de clasificación del dato para gestión de su ciclo de vida. ▪ Requerimientos asociados al autoservicio del dato consecuencia de la clasificación.
--	---

ACTIVIDADES

1. Arquitectura del dato

Clasificar el dato según su naturaleza y metadata. Si existen datos que contemplen distintas naturalezas, ya que puede provenir de distintas fuentes, se debe crear la relación y almacenar dichas relaciones.

2. Ordenar el dato a través criterios clasificados tales como tiempo de creación o modificación, tiempo de último acceso o publicación, predecesores, alfabético por nombre, según flujo de procesos, etc. Los datos pueden llegar de formas aleatorias y se deben ordenar para maximizar su oportunidad y uso. Aplicar a la clasificación la normativa proveniente de la ley de transparencia y acceso a la información que define los procedimientos para el ejercicio y garantía de acceso a la información y establece las excepciones a la publicidad de la misma. Así mismo, aplicar los lineamientos del Marco de Referencia de AE relacionados con taxonomía y diccionario de datos para gestión del ciclo de vida.

3. Calidad del dato.

4. Catalogar los criterios y campos críticos para calidad del dato. Catalogar los parámetros para depuración el dato.
5. Enmascaramiento de datos. Catalogar campos para enmascaramiento del dato.
6. Reducción (entornos o datos). Catalogar tipos para reducción del dato y sus entornos.
7. Archivado de datos. Catalogar plataformas y herramientas que implican conservación del dato.
8. Autoservicio del dato
 - a. Parametrización del modelo de datos de las aplicaciones y conservar esta catalogación.
 - b. Definición de conceptos de la actividad misional y casos de prueba.
 - c. Catalogar las operaciones que hacen predictivo al dato.
 - d. Catalogar los criterios de oportunidad, valor, y uso en donde se alinea el dato con procesos de negocio.
 - e. Catalogar las audiencias involucradas para el contenido del dato.
 - f. Catalogar los niveles de servicio por demanda para el dato.
 - g. Catalogar los requerimientos de infraestructura por demanda para conservar el dato.

Tabla 3. Paso 3: Clasificar

7.4 Paso 4: Analizar.

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 4.

PASO4: Analizar

<p>OBJETIVOS</p>	<ul style="list-style-type: none"> ▪ Analizar el dato, previo uso por consumidores, en los ejes del ciclo de vida: enmascaramiento, reducción, archivado, generación, y autoservicio. ▪ Establecer los requerimientos finales que definen el ciclo de vida del dato considerado..
<p>ENTRADAS</p>	<p>SALIDAS</p>
<ul style="list-style-type: none"> ▪ Dato clasificado por metadata y con orientación al autoservicio. ▪ Referencia única de clasificación del dato para gestión de su ciclo de vida. ▪ Requerimientos asociados al autoservicio del dato consecuencia de la clasificación. 	<ul style="list-style-type: none"> ▪ Reporte final de ejes del ciclo de vida del dato considerado. ▪ Anexo técnico con requerimientos específicos para ciclo de vida del dato considerado.
<p>ACTIVIDADES</p>	
<ol style="list-style-type: none"> 1. Arquitectura del dato <ol style="list-style-type: none"> a. 1.1. Realizar las interpretaciones a la metadata del dato para acotar las posibles interpretaciones que el usuario realiza a la forma como se estructura el dato, por ejemplo, el formato de la fecha o los campos obligatorios. b. 1.2. Definir si el dato posee los roles y las responsabilidades para custodia de la información del dato. 2. Calidad del dato 3. Analizar los criterios, campos críticos, y necesidades de depuración pertinentes en calidad para gestión del ciclo de vida. 	

4. Enmascaramiento de datos
5. Analizar las estrategias pertinentes para enmascaramiento del dato.
6. Reducción (entornos o datos)
7. Analizar las estrategias existentes para reducción del dato y sus entornos.
8. Archivado de datos
9. Analizar las estrategias definidas para conservación del dato en sus varios entornos.
10. Autoservicio del dato
11. Analizar relaciones del dato con otros aplicativos.
12. Revisar el manejo de la solicitud de extracción de datos en producción para el dato.
13. Analizar las audiencias involucradas para el contenido del dato y la normativa de acceso.
14. Analizar la demanda de servicios alrededor del dato.
15. Analizar la demanda para cumplir con la conservación el dato.
16. Reporte del dato
17. Generar reporte final de análisis del dato para gestionar su ciclo de vida efectivamente.
18. Desarrollar anexo técnico con los requerimientos que se avalan para gestionar efectivamente el ciclo de vida del dato.

Tabla 4. Paso 4: Integrar MDM

7.5 Paso 5: Rediseñar los procesos

A continuación, se describen los objetivos, entradas, salidas y actividades del paso

5.

PAS05: Rediseñar los procesos

<p>OBJETIVOS</p>	<ul style="list-style-type: none"> ■ Definir operaciones aplicables al dato con respecto a los ejes del ciclo de vida aceptados. ■ Realizar modificaciones sobre sistemas, entornos, e infraestructura de información resultado de los requerimientos finales para ciclo de vida del dato.
<p>ENTRADAS</p>	<p>SALIDAS</p>
<ul style="list-style-type: none"> ■ Reporte final de ejes del ciclo de vida del dato considerado. ■ Anexo técnico con requerimientos específicos para ciclo de vida del dato considerado. ■ Políticas de la institución para sistemas, entornos, e infraestructura. 	<ul style="list-style-type: none"> ■ Definición de cada operación aplicable al dato en su gestión del ciclo de vida. ■ Ajuste de políticas de la institución para considerar requerimientos de ciclo de vida del dato en sistemas, entornos, e infraestructura. ■ Ajuste de operaciones sobre sistemas, entornos, e infraestructura para implementar requerimientos de ciclo de vida del dato.
<p>ACTIVIDADES</p>	
<p>1. Arquitectura del dato</p> <p>1.1. Modificar la definición del dato a partir de las posibles operaciones que se realizan sobre este. Estas operaciones se pueden realizar sobre los datos cuantitativos: sumas, restas, porcentajes, y los no cuantitativos: conteos por reglas de negocio.</p>	

<ul style="list-style-type: none"> 1.2. Definir características mínimas de presentación de los datos, tales como: formato, nivel de acceso, difuminación de valor, entre otros. 1.3. Definir reglas de negocio que permitan dar significancia a los datos, a través de agrupaciones de tipo funcional, y características de publicación.
<ul style="list-style-type: none"> 2. Calidad del dato <ul style="list-style-type: none"> 2.1. Implementar tareas para aseguramiento y depuración de calidad durante el ciclo de vida del dato.
<ul style="list-style-type: none"> 3. Enmascaramiento de datos <ul style="list-style-type: none"> 3.1. Implementar estrategias y herramientas para tratamiento del enmascaramiento del dato.
<ul style="list-style-type: none"> 4. Reducción (entornos o datos) <ul style="list-style-type: none"> 4.1. Implementar los tipos de entorno y las variantes del dato resultado de la reducción para el ciclo de vida.
<ul style="list-style-type: none"> 5. Archivado de datos <ul style="list-style-type: none"> 5.1. Implementar los mecanismos en plataformas y herramientas para realizar la conservación del dato.
<ul style="list-style-type: none"> 6. Autoservicio del dato <ul style="list-style-type: none"> 6.1. Enmascaramiento coherente de datos sensibles, teniendo en cuenta políticas de seguridad, privacidad, y datos abiertos. 6.2. Descarga y enmascaramiento de datos coherentes de producción en entorno de pruebas. 6.3. Implementar plataformas y herramientas para demanda del dato en un contexto de autoservicio.
<ul style="list-style-type: none"> 7. Acciones del dato <ul style="list-style-type: none"> 7.1. Implementar las operaciones requeridas sobre sistemas, entornos, e infraestructura para implementar requerimientos de ciclo de vida del dato.

Tabla 5. Paso 5: Manipular

7.6 Paso 6: Consolidar

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 6.

PASO 6: Consolidar	
OBJETIVOS	<ul style="list-style-type: none"> Consolidar los ambientes técnicos impactados por el ciclo de vida del dato considerado. Abordar los requerimientos para anonimización del dato en su ciclo de vida.
ENTRADAS	SALIDAS
<ul style="list-style-type: none"> Normativa para anonimización de datos en sector público. Ajuste de políticas de la institución para considerar requerimientos de ciclo de vida del dato en sistemas, entornos, e infraestructura. Medición de operaciones sobre sistemas, entornos, e infraestructura para soportar requerimientos de ciclo de vida del dato. 	<ul style="list-style-type: none"> Reporte de consolidación de entornos para ciclo de vida del dato. Definición de enmascaramiento del dato por entorno adoptado para ciclo de vida. Implementación de requerimientos en entornos adoptados para el ciclo de vida del dato.
ACTIVIDADES	

<ol style="list-style-type: none"> 1. Enmascaramiento de datos Optimizar alternativas para tratamiento del enmascaramiento del dato. 2. Reducción (entornos o datos) Reducción coherente del volumen de entornos no productivos en donde se gestiona el ciclo de vida del dato. Incluye generación de datos para prueba. 3. Consulta de datos extraídos y solicitud de carga en entorno definido. <ol style="list-style-type: none"> 3.1. Carga de datos en entorno seleccionado teniendo en cuenta la reducción de campos, tipos, y valores. 3.2. Archivado de datos 3.3. Consolidación y virtualización de los mecanismos para realizar la conservación del dato. 4. Autoservicio del dato Consolidación de las opciones de demanda del dato disponibles para gestión del ciclo de vida.
--

Tabla 6. Paso 6: Manipular

7.7 Paso 7: Conservar

A continuación, se describen los objetivos, entradas, salidas y actividades del paso 7.

PASO 7: Conservar	
OBJETIVOS	<ul style="list-style-type: none"> ■ Almacenar la información asociada al dato según los requerimientos de la gestión de su ciclo de vida. ■ Reportar información resultado del autoservicio del dato en repositorios y sitios para

	<p>evaluación de uso, creación de valor, y mantenimiento de trazabilidad.</p>
ENTRADAS	SALIDAS
<ul style="list-style-type: none"> ■ Reporte de consolidación de entornos para ciclo de vida del dato. ■ Definición de enmascaramiento del dato por entorno adoptado para ciclo de vida. ■ Implementación de requerimientos en entornos adoptados para el ciclo de vida del dato. 	<ul style="list-style-type: none"> ■ Reportes de almacenamiento y conservación periódica de información del dato en entornos e infraestructura. ■ Pruebas de existencia de la conservación del dato para continuidad, replicación, análisis histórico, etc.
ACTIVIDADES	
<ol style="list-style-type: none"> 1. Enmascaramiento de datos Conservación de opciones de enmascaramiento para el dato. 2. Reducción (entornos o datos) <ol style="list-style-type: none"> 2.1. Conservación de configuración de entornos no productivos en donde se gestiona el ciclo de vida del dato. 2.2. Conservación de la reducción de campos, tipos, y valores del dato por entorno seleccionado. 3. Archivado de datos <ol style="list-style-type: none"> 3.1. Almacenar los datos en su forma atómica: nivel de granularidad mínima. 3.2. Almacenar las transformaciones aplicadas sobre el dato. 3.3. Almacenar los datos bajo clasificaciones y agrupaciones realizadas. 4. Autoservicio del dato 	

Consolidación de las opciones de demanda del dato disponibles para gestión del ciclo de vida

Tabla 7. Paso 7: Conservar

Nota. Los pasos anteriormente descritos contienen solo las disposiciones genéricas o estándares para un dato sin tener en cuenta su campo de aplicación/estudio. Para el caso de datos especializados como por ejemplo: Manejo de audio, de video, de georreferenciación, criptográfico, etc., las consideraciones pueden contemplar la omisión de alguno de estos pasos o la inclusión de nuevos pasos.

8 GLOSARIO

Ciclo de vida del dato: Es una práctica basada en políticas que se encarga del flujo de los datos de los sistemas de información a través de su ciclo de vida: desde la creación y el almacenamiento inicial, hasta el momento cuando se convierte en obsoleto y es eliminado.

Datos: Los datos son las piezas individuales o recolección de hechos, cantidades, caracteres, símbolos y en general elementos crudos de conocimiento que pueden ser persistidos y relacionados de alguna manera por la institución, ya sea en medio físico o electrónico, y que no es necesario que hayan tenido un procesamiento, cálculos o estructuras elaboradas previas en su proceso de construcción.

Enmascaramiento: Es un proceso que consiste en transformar los datos a partir de la aplicación de algoritmos de una sola vía, que permita a los sistemas procesar los datos, pero que a su vez no habilite a los usuarios finales a conocer los datos originales, los datos enmascarados deben conservar el mismo significado y sentido que los datos originales.

RAW DATA: Se refiere a datos recolectados de determinada fuente, que aún no han sido sujeto de transformaciones ni enmascaramiento, también se les llamo datos primarios.

Rollback: Mecanismo de “marcha hacia atrás” que permite deshacer las operaciones efectuadas en una transacción.

9 REFERENCIAS

[1] M. Mosley. Functional Framework. The Data Management Association. DAMA-Data Management Body of Knowledge - DMBOK. Versión 3.02. (2008).

