

Metodología  
y lineamientos  
sobre el ciclo  
de vida de la  
**Gestión de  
Información**



**Oficina de Gestión de la Información**  
Equipo de Gestión de Información Sectorial

## Contenido

1. CONTEXTO.....	4
2. CICLO DE VIDA DE LA GESTIÓN DE LA INFORMACIÓN.....	6
2.1 Crear, Capturar y Recopilar .....	7
2.2 Organizar y Almacenar .....	8
2.3 Analizar y Utilizar .....	10
2.4 Compartir .....	11
2.5 Reutilización y mantenimiento .....	13
2.6. Archivar / destruir .....	14
3. TEMAS TRANSVERSALES AL CICLO DE VIDA.....	15
3.1 Calidad de los datos .....	15
3.2 Integración de datos .....	16
3.3 Seguridad de los datos .....	17

## Ilustraciones

Ilustración 1 – Componentes del SIMEC.....	6
Ilustración 2. Hitos del Plan .....	8
Ilustración 3 – Avances 2021 y Metas 2022 en función de los Hitos de la Gestión de Información .....	13
Ilustración 4 – Componentes de la Gestión del Portafolio de Proyectos.....	20

## Tablas

Tabla 1. Lineamientos por temas .....	21
---------------------------------------	----

## 1. CONTEXTO

Como parte del Modelo de Gestión de Información del sector minero energético se ha propuesto la identificación y consolidación de los activos de información del sector, con el fin de garantizar su calidad, disposición, acceso y uso, para promover la transparencia, el cumplimiento de los objetivos sectoriales y una relación más consolidada con los usuarios y los ciudadanos.

Siguiendo los lineamientos del Ministerio de las Tecnologías de Información y Comunicaciones (MINTIC) en cuanto al gobierno digital y el Marco 2.0 para la Arquitectura Empresarial, así como los avances sectoriales en materia de gobierno de datos, se propone la definición del ciclo de vida de la gestión de la información.

Este tema es relevante en la medida en que contribuye a identificar las etapas en las que se produce, almacena, utiliza y distribuye la información, evidenciando la necesidad de abordar temas tan indispensables como la definición de criterios de interoperabilidad e integración, lineamientos de calidad y seguridad, entre otros.

El beneficio de la gestión del ciclo de vida de la información es que el proceso puede ayudar a las entidades a gestionar roles, responsabilidades y obligaciones a lo largo de la existencia de cualquier activo de información en particular.

La propuesta evidencia la necesidad de articular los componentes técnicos del ciclo de vida, con políticas, estándares y procedimientos, e incluso normativas sobre la privacidad y los niveles de acceso a los datos.

Tener una estrategia sólida de gestión de información que defina e implemente las etapas del ciclo de vida de los activos de información permite aprovechar al máximo los esfuerzos e inversiones que han realizado las instituciones del sector, en la producción, administración y publicación de los activos de información. Así mismo, contribuye a que las entidades tengan un mayor control sobre sus datos, ayuda a satisfacer las necesidades de archivo, minimiza la carga de TI, reduce los costos de almacenamiento, permite una toma de decisiones más acertada y una recuperación rápida durante una crisis.

A partir de los avances obtenidos en 2021 por las instituciones adscritas, en cuanto a los inventarios de activos de información y los catálogos de componentes, se plantea para 2022, aplicar el Modelo de Gestión de Información Sectorial, con la identificación de cómo se relacionan dichos activos y componentes, con etapas del ciclo de vida.

Se destaca que este tema está muy asociado al Pilar de Gobierno de Datos, sin embargo, es necesario proponer una estrategia más integral, que incluya no solo los datos maestros identificados, sino la aplicación de los parámetros del ciclo de vida a la información

geográfica y estadística, la implementación de la Infraestructura de Datos Espaciales, y en general a la definición de los datos fundamentales del sector y la estrategia del SIMEC.

Adicionalmente, instituciones como la UPME, el SGC y la ANM han avanzado en sus proyectos de Gobierno de Datos y en las definiciones sobre políticas, metadatos y datos maestros, entre otros. En tal sentido, se plantea una valiosa oportunidad de colaborar y coordinar acciones sobre las etapas del ciclo de vida de la gestión de la información, de manera que se articulen los temas comunes del gobierno de datos, la IDE y la Arquitectura Empresarial, principalmente,

## 2. CICLO DE VIDA DE LA GESTIÓN DE LA INFORMACIÓN

La aplicación de lineamientos y mejores prácticas sobre gestión de información es esencial para el uso eficaz y eficiente de los datos a lo largo del ciclo de vida de los activos.

La gestión del ciclo de vida implica comprender cómo se debe crear, gestionar, utilizar y eliminar la información. La gestión del ciclo de vida tiene como objetivo reducir las ineficiencias y garantizar el mantenimiento y la protección de la información que tiene valor a largo plazo para las instituciones, el sector, el gobierno y los usuarios.

Cada fase del ciclo de vida es fundamental para recopilar, gestionar y mantener conjuntos de datos útiles y de alta calidad. Además de ser un proceso cíclico, cada fase puede y, en ocasiones, debe ser iterativa.

El ciclo de vida de los datos consta de las siguientes fases:



Las características centrales adicionales al ciclo de vida que se aplican en cada etapa son calidad, integración y seguridad.

El ciclo de vida de los datos como un proceso estandarizado, en el marco de la gestión de información, se gestiona apropiadamente cuando se aborda desde la generación de la información y el diseño de los servicios, hasta la disposición final. Esto incluye aspectos legales como la protección de datos personales, los niveles de acceso, custodia y

responsabilidad en su administración y la definición de cuándo y cómo se preservan o se destruyen los datos al final.

A continuación, se plantean los principales aspectos que se deben tener en cuenta en las etapas del ciclo de vida de la gestión de la información del sector minero energético:

## 2.1 Crear, Capturar y Recopilar

El ciclo de vida de la gestión de datos comienza con la planificación de la creación, recopilación, captura o adquisición de datos. También es el punto de entrada o la etapa en la cual se consolidan los activos de información que provienen de otras fuentes y procesos.

Esta etapa del ciclo de vida incluye las siguientes actividades:

### Definir requerimientos

Con el fin de establecer el enfoque adecuado para el sector, se requiere determinar las necesidades y requerimientos necesarios para respaldar el uso y la gestión de la información como un activo a lo largo de su vida útil.

Los requerimientos de información del activo deben definirse a partir del análisis de los datos que se necesitan para la toma de decisiones estratégicas y para el cumplimiento de los objetivos institucionales y sectoriales.

Los requerimientos de información, no necesariamente se conocen o se planifican desde el principio, sin embargo, un enfoque estructurado, ayudará a que se complementen las definiciones en las demás etapas del ciclo de vida, en la medida en que se avanza en la gestión de la información.

### Minimizar la recopilación de datos

Es importante definir los requerimientos de información realistas y específicos y evitar pedir todo "por si acaso". La recopilación de más datos de los necesarios, puede crear una carga de procesamiento y almacenamiento adicional, y aumentar los riesgos de privacidad y seguridad. Por lo tanto, es importante considerar los costos de recopilar y mantener dichos datos frente a los beneficios que proporcionarán.

También puede ocurrir que los datos requeridos ya existan. Por tal razón, antes de proponer la recopilación o generación de nuevos datos, es importante verificar que no estén disponibles en otras instituciones o en el portal de datos abiertos del gobierno colombiano.

### Determinar el método de recolección más apropiado

Se debe evitar, en lo posible, la recolección manual de los datos, en especial en la digitalización de información que proviene de formularios u otros documentos análogos.

En la actividad de recolección se deben seleccionar las tecnologías más apropiadas como el uso de sensores, imágenes de sensores remotos, escáneres laser, técnicas de fotogrametría, etc., con el fin de optimizar y agilizar los procesos de recopilación de datos.

En este punto es importante considerar tecnologías inteligentes e internet de las cosas, como instrumentos avanzados de captura de datos e información.

## 2.2 Organizar y Almacenar

Para respaldar la gestión de activos de información, se deben organizar y almacenar grandes cantidades de datos de manera adecuada y eficiente. La elección entre las opciones de almacenamiento depende de muchos factores y nunca es única.

Los entornos de almacenamiento deben cumplir con los requisitos de información organizacional (OIR), cumplir con la legislación y las políticas pertinentes, ser interoperables entre los diferentes sistemas de gestión de la información y permitir que los datos se almacenen y gestionen durante la vida útil del activo.

La selección de almacenamiento se verá afectada por consideraciones específicas, que incluyen:

- Uso previsto de los datos (por quién y con qué propósito)
- Características de la base de datos o los sistemas de información existentes.
- Tipo y volumen de los datos a integrar
- La frecuencia con la que se accederá a los datos
- La frecuencia con la que se actualizarán los datos.
- La velocidad a la que será necesario acceder a los datos, especialmente en situaciones de emergencia.
- Tecnología actualmente disponible

Es importante mantener un entorno común para el almacenamiento de datos. Si bien los datos pueden estar alojados en varios sistemas de almacenamiento o administración, un solo portal puede proporcionar un único punto de acceso que reúne esos datos para su visualización y el análisis.

La definición de una estrategia de almacenamiento y administración ayuda a evitar la duplicación que resulta de almacenar los mismos datos en diferentes ubicaciones. Además del trabajo adicional necesario para actualizar los datos en varias ubicaciones, la duplicación aumenta el riesgo de que los datos se modifiquen en una ubicación, pero no en otras, lo que aumenta el riesgo de que se utilice información incorrecta o desactualizada.

A continuación, se plantean los temas claves relacionados con la organización y mantenimiento de información

### Diseñar para la interoperabilidad

La interoperabilidad técnica se refiere a la capacidad de diferentes productos o sistemas de diferentes proveedores de servicios para intercambiar información entre sí, de manera que puedan trabajar juntos sin problemas, ya sea en el presente o en el futuro.

La interoperabilidad implica el uso de estándares, protocolos de comunicación y tecnologías que pueden ser muy diferentes, pero que pueden comunicarse entre sí. En este tema, se siguen los lineamientos y requisitos del MINTIC, quien recomienda seleccionar tecnologías abiertas y / o plataformas independientes del proveedor cuando estén disponibles y se adapten a las necesidades de las instituciones.

#### Establecer un entorno de almacenamiento seguro

El nivel de seguridad aplicado al almacenamiento de datos debe estar alineado con la clasificación de sensibilidad y seguridad de los datos. Los activos de información pueden contener información confidencial o clasificada de seguridad, incluida información legal y financiera.

En el marco del Modelos de Gestión de Información del sector, se deben definir y aplicar los lineamientos de política en cuanto a la seguridad y privacidad de la información, como un punto de partida para definir de manera adecuada, las opciones de almacenamiento de la información.

Es importante destacar que se deben cumplir los requisitos legales tales como el manejo de los datos personales, lo cual, también implica dar especial atención a este tipo de información antes de definir las soluciones de almacenamiento y administración

#### Garantizar el acceso a los datos

La accesibilidad de los datos de los activos no solo es una práctica de gestión importante, sino que también está definida en la Guía de Datos Abiertos<sup>1</sup> del MINTIC, en la cual se define que, según la Ley 1712 de 2014 sobre Transparencia y Acceso a la Información Pública Nacional, los datos abiertos son “todos aquellos datos primarios o sin procesar, que se encuentran en formatos estándar e interoperables que facilitan su acceso y reutilización, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos”. De este modo, la Ley establece la obligatoriedad de las entidades públicas de “divulgar datos abiertos”, teniendo en cuenta las excepciones de acceso a la información, asociadas a información clasificada y reservada establecidas en su título tercero, artículos 18 y 191.

Lograr que los datos sean de libre acceso permite que otras organizaciones y el público se beneficien e innoven utilizando los datos generados. Sin embargo, se reconoce que los riesgos asociados al tema de seguridad deben equilibrarse frente a las estrategias de datos abiertos, lo cual se debe considerar y definir de manera coordinada con las instituciones adscritas al sector minero energético, debido a que indispensable alinear los permisos basados en roles o basados en identidad con las clasificaciones de los datos, y reconocer que no todos los datos relacionados con un activo pueden tener la misma clasificación y pueden requerir la administración de varios niveles de permisos diferentes.

---

1

<https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf>

## 2.3 Analizar y Utilizar

El análisis de datos es una parte clave del Modelo de Gestión de Información Sectorial porque permite transformar grandes cantidades de datos en información útil.

La analítica puede ayudar a identificar y resolver problemas o predecir problemas que aún no han ocurrido. Por ejemplo, la analítica se puede utilizar para optimizar la gestión porque permite:

- Definir programas de mantenimiento más eficaces y eficientes
- Monitorear la condición y el desempeño de los activos
- Identificar brechas en la implementación de los pilares del modelo
- Asegurar la planificación y desarrollo de las capacidades institucionales y sectoriales

Los aspectos más importantes a considerar en esta etapa del ciclo de vida de la gestión de la información se proponen a continuación:

### Análisis de datos

El análisis de datos debe ser adecuado para su propósito. El tipo de análisis utilizado dependerá de la pregunta que se esté investigando. El análisis de datos utilizado puede ser reactivo o proactivo; estas categorías se encuentran en un continuo desarrollo, desde la retrospectiva hasta la prospectiva, de acuerdo con las siguientes definiciones básicas:

### Análisis Reactivo

El análisis reactivo puede ayudar a las organizaciones a responder rápidamente a problemas operativos o requisitos de mantenimiento.

- **Descriptivo:** esta forma de análisis se aplica para comprender qué sucedió con el activo.
- por ejemplo, pérdidas pasadas, permite la previsión de pérdidas e identifica la causa de un incidente.
- **Diagnóstico:** esta forma de análisis se aplica para averiguar por qué está sucediendo, por ejemplo, qué proceso y condiciones están creando la situación.

### Análisis Proactivo

El análisis proactivo puede ayudar a las organizaciones a planificar mejor las necesidades futuras de los usuarios.

- **Predictivo:** esta forma de análisis se aplica para predecir lo que sucederá con el activo y a menudo utiliza algoritmos de aprendizaje automático.
- **Prescriptivo:** esta forma de análisis de datos se aplica para comprender las mejores acciones que se puede tomar en una situación particular para cambiar la operación del proceso.

### Conceptos clave de análisis de datos

**Elaboración de perfiles de datos:** Comprender el perfil de los datos incluye la evaluación de la calidad y la identificación de problemas potenciales con los datos, incluida la comparación de formatos de datos con estándares para garantizar la interoperabilidad, identificando las diferencias de las

expectativas dentro de los conjuntos de datos. Se pueden utilizar herramientas automatizadas para optimizar estos procesos.

**Limpieza de datos:** Uno de los primeros pasos del análisis de datos (después de definir los requisitos de datos, recopilar y almacenar los datos) es limpiarlos. La limpieza de datos es vital para obtener datos precisos (los datos incorrectos pueden generar resultados engañosos). El análisis de datos y el uso de técnicas para automatizar estos métodos de verificación de errores pueden ayudar a acelerar este proceso; sin embargo, un analista de datos aún debe participar para investigar cualquier problema.

**Ingeniería y modelado de datos:** Un analista a menudo necesita combinar conjuntos de datos y construir modelos con múltiples capas de datos para generar conocimientos de datos. El modelado de datos es cuando un científico de datos crea un modelo de datos para correlacionar los datos, a menudo con los resultados comerciales en mente.

**Visualización y comunicación de datos:** La comunicación es el último paso del proceso de análisis de datos y, a menudo, se pasa por alto. Los datos deben entregarse a la organización de manera significativa para respaldar la toma de decisiones. La visualización de datos se trata de la representación visual de datos como medio de comunicación. Por ejemplo, una forma común de visualización es a través de sistemas geoespaciales y los gemelos digitales (digital twins<sup>2</sup>).

**Uso de datos para cumplir objetivos específicos:** Los datos generados o recopilados se pueden usar para rastrear y medir continuamente muchos factores, incluido el uso, la condición y las excepciones. Sin una definición clara de lo que se busca en cada etapa del ciclo de vida de los activos de información, el progreso puede retrasarse, la recopilación y el uso de datos pueden volverse costosos y no enfocados, y el valor de los activos de información puede disminuir con el tiempo.

Una vez que se entregan los objetivos bien definidos, se puede explorar el aprovechamiento de los datos para identificar nuevas formas de operar y nuevas innovaciones en los servicios. Al combinar varios conjuntos de datos y aprender de otros proyectos, los datos se pueden usar para desarrollar nuevos activos e innovaciones para lograr una gestión más inteligente.

En el proceso de implementación del Modelo de Gestión de Información, se deben considerar cómo los activos de información se pueden aprovechar de manera más efectiva en el sector o estar disponibles para la innovación como un recurso de datos abiertos.

## 2.4 Compartir

Al trabajar con datos, es importante recordar que los datos se pueden cerrar, compartir o abrir debido a la sensibilidad, el nivel de riesgo y los lineamientos de uso y publicación. Al comprender de dónde provienen los datos, quién puede usarlos y qué se puede hacer con ellos, se pueden optimizar las oportunidades de compartir y usar los activos de información.

---

<sup>2</sup> Digital twins: réplica digital de un producto, servicio o proceso. La idea es someter dicho producto o servicio a estrés de manera que se vean testadas sus principales debilidades sin la necesidad de construir costosos prototipos. <https://www.aecoc.es/innovation-hub-noticias/que-son-los-digital-twins-y-por-que-son-tan-importantes/>

El intercambio de datos es un requisito fundamental para la gestión de información, debido a que, en la mayoría de los proyectos, existen instituciones y partes interesadas que requieren servicios a lo largo del ciclo de vida de los activos de información.

La necesidad de intercambiar datos entre las instituciones y partes interesadas, de manera oportuna y eficiente es clave para el éxito de la gestión de activos de información, la creación de servicios ciudadanos digitales y la consolidación de la plataforma de integración sectorial, entre otras aplicaciones.

Es importante resaltar que se debe promover la publicación de datos e información, siempre y cuando se tenga un manejo consciente de los temas de seguridad y privacidad, y existan las protecciones adecuadas, de acuerdo con los lineamientos de política definidos.

Entre los principales beneficios de compartir los datos se encuentran:

- Mejor entendimiento de la existencia y disponibilidad de información en el sector
- Toma de decisiones acertada, con base en información confiable y disponible
- Incremento en la productividad para el cumplimiento de los objetivos sectoriales
- Procesos de planeación del desarrollo del sector eficientes

Con el fin de garantizar la formulación de una adecuada estrategia de publicación de información, se deben considerar los siguientes aspectos:

- Definiciones sobre privacidad, en particular lo relacionado con protección de datos personales
- Lineamientos sobre propiedad intelectual, confidencialidad y responsabilidad
- Seguridad de Información
- Gobierno de datos, calidad e interoperabilidad

En cuanto a los tipos de datos que se pueden compartir, se propone tener en cuenta las siguientes consideraciones:

#### Datos cerrados o seguros

Los datos cerrados o seguros son datos que solo las personas dentro de una organización pueden ver y usar. Se deben identificar las razones por las cuales los datos están cerrados y por qué no deberían estar disponibles en el dominio público. Los datos cerrados generalmente se asocian con operaciones sensibles o críticas.

#### Datos compartidos

Los datos compartidos son datos que se comparten con una organización específica, o grupo de organizaciones o personas, para un propósito específico. El intercambio de datos es la forma en que las instituciones pueden proporcionar acceso autorizado a los datos que tienen de manera controlada, para generar mejores resultados a los ciudadanos. En este contexto, se plantean los siguientes principios para el intercambio de datos, de manera segura:

- Compartir datos para propósitos apropiados y autorizados
- Compartir datos solo con usuarios autorizados
- Usar datos en un ambiente seguro y protegido
- Aplicar protecciones apropiadas a los datos

- Asegurar que los resultados públicos de los proyectos de intercambio de datos no identifiquen a las personas

Si las protecciones conjuntas que ofrecen los Principios no son suficientes para proteger contra el riesgo de violación de datos o reidentificación de datos, entonces los datos no deben compartirse.

### Datos abiertos

Los datos abiertos son datos a los que cualquiera puede acceder, usar y compartir. El gobierno nacional ha abierto el acceso a diversos conjuntos de datos, lo cual permite que los ciudadanos utilicen estos datos y descubran nuevas soluciones para el beneficio de todos.

Sin embargo, la simple divulgación de datos no es suficiente y para que los datos se consideren verdaderamente abiertos, los propietarios deben indicar claramente que otras organizaciones o personas pueden usarlos de la forma que deseen, ya que, sin un permiso expreso, los datos no pueden considerarse abiertos.

De acuerdo con la guía de datos abiertos del MINTIC, se recomienda consultar el instructivo sobre licencias de datos<sup>3</sup> elaborado por IDECA (Infraestructura de Datos Espaciales del Distrito Capital).

En esta misma guía se plantean los principios de los datos abiertos:

- Abiertos por defecto: la información debe estar disponible para consulta ciudadana desde el momento de su generación
- Oportunos y completos: los datos abiertos son relevantes sólo si aportan valor para el usuario de la información
- Accesibles y utilizables: Son gratuitos y bajo licencia abierta, facilitando los datos a la mayor cantidad de usuarios y máquinas
- Comparables e interoperables: Pueden ser comparados desde distintos sectores, territorios y periodos de tiempo
- Mejoran la gobernanza y la participación ciudadana: abrir datos fortalece la gobernanza y favorece la transparencia
- Apoyan el desarrollo incluyente y la innovación: el uso de datos abiertos permite la construcción de nuevo conocimiento que puede ser usado en la generación de valor social y económico

## 2.5 Reutilización y mantenimiento

La reutilización de los activos de información del sector minero energético es un aspecto determinante en la creación de valor para las instituciones adscritas, la industria y los ciudadanos, en general. En el contexto de la transformación digital, la reutilización permitirá romper los siles y barreras existentes en cuanto al uso de la información.

La reutilización es posible gracias a la implementación de estándares, sistemas de datos interoperables, formatos de datos comunes, el intercambio proactivo y la liberación de datos abiertos.

---

<sup>3</sup> <https://www.ideca.gov.co/sites/default/files/documentacion/instructivolicensadatos.pdf>

De otra parte, el mantenimiento continuo de los activos de información es un proceso fundamental, en particular cuando se manejan grandes volúmenes de datos y se proyecta la implementación de componentes de BIG DATA como el Hub de información sectorial, la interoperabilidad de sistemas de información y los modelos de analítica avanzada para el sector. Así, es necesario contar con procesos para monitorear, mantener y actualizar los datos para garantizar la eficiencia y la mejora continuas, teniendo en cuenta temas tan importantes como los siguientes:

#### Monitoreo continuo de los activos de datos

Se requiere hacer un monitoreo continuo de la efectividad de la recopilación y consolidación de los activos de información, así como del uso y la gestión de los requisitos cambiantes durante todo el ciclo de vida.

Los cambios podrían ser en los activos en sí mismos, cómo se administran, la tecnología que los respalda o los objetivos sectoriales que se quieren lograr. Los propietarios de activos deben reevaluar los enfoques de gobernanza de datos con regularidad para gestionar los cambios, la privacidad y la seguridad. Esto asegurará que los datos estén actualizados cuando las situaciones cambien, sin volverse redundantes y engañosos.

Los propietarios de los datos son responsables de garantizar que los activos se mantengan actualizados, mediante un programa de mantenimiento, el cual debe se debe implementar, monitorear e incluir procesos claros y puntos de responsabilidad para mantener y actualizar los datos en cada fase del ciclo de vida del activo.

#### Validación de datos

Para mantener el nivel requerido de calidad de los datos, es importante llevar a cabo la validación de datos en función de los requisitos identificados en la evaluación de las necesidades de datos. Siempre que sea posible, la validación de datos debe automatizarse y los errores deben corregirse en la fuente. La validación de datos también debe incluir la detección y mitigación de datos maliciosos. La validación de contenido malicioso es importante para proteger no solo la integridad de los datos sino también el sistema. Los datos maliciosos que no están bloqueados por la validación de datos se pueden usar para explotar vulnerabilidades dentro de un sistema y crear un impacto en otros componentes de un sistema, así como en otros usuarios.

#### Consideraciones de seguridad

Ya sea que los datos estén en reposo, en uso o en tránsito, se deben realizar controles de seguridad adecuados para protegerlos. Los requisitos de seguridad de los datos y los sistemas pueden cambiar durante el ciclo de vida del proyecto, ya que los datos centrales se actualizan con nuevos datos a lo largo del tiempo y el perfil de riesgo de los datos puede cambiar. Por lo tanto, es fundamental evaluar los cambios de riesgo y los requisitos de seguridad a lo largo del ciclo de vida. Al evaluar y reevaluar, los riesgos residuales, así como los riesgos recientemente identificados, pueden mitigarse con los controles adecuados.

## **2.6. Archivar / destruir**

Cuando un activo de información llega al final de su vida, se debe archivar o destruir. Para evitar grandes costos de almacenamiento (o minimizar el riesgo de destrucción prematura de datos), se

deben evaluar e identificar las consideraciones que se aplican a la retención y destrucción de datos, de acuerdo con las normas y requisitos legales en temas como protección de datos personales, lineamientos del Archivo General de la Nación y otras disposiciones del gobierno.

Cualquier decisión de archivar o destruir datos también debe tomarse de acuerdo con los registros de las instituciones y los requisitos de gestión de la información. Todas las decisiones de retención y destrucción deben estar autorizadas y documentadas para lograr transparencia y responsabilidad sobre la destrucción de los activos de información del sector. Si en los procesos de gestión documental, se trabaja con proveedores externos, se debe garantizar que ellos puedan respaldar e implementar los marcos de retención y destrucción de datos requeridos.

### 3. TEMAS TRANSVERSALES AL CICLO DE VIDA

Como complemento a las etapas del ciclo de vida de la información se deben abordar los temas de calidad, integración y seguridad.

#### 3.1 Calidad de los datos

Una vez que se han identificado los requisitos de información y los estándares relevantes, se debe decidir sobre la calidad de los datos, la cual debe garantizar que los datos sean adecuados para su propósito.

De acuerdo con el Modelo DAMA, las dimensiones de calidad de datos definidas incluyen: Exactitud, Integridad, Coherencia, Razonabilidad, Actualidad, Unicidad / Duplicación, Validez y Accesibilidad.

Sin un nivel suficiente de confianza en los datos, es imposible desarrollar una visión precisa de los procesos y los requerimientos de información, lo cual puede conducir a una mala toma de decisiones.

Como recomendación general, se prefiere la calidad que la cantidad de los datos. Es mejor tener un conjunto de datos bien estructurado que haya sido verificado y validado, y que sea confiable, que un gran volumen de datos poco confiables y mal organizados.

#### Requisitos de calidad de los datos

Los problemas de calidad de los datos causados por averías o calibraciones de los dispositivos de captura pueden generar datos incorrectos o inexactos. Si esto representa un riesgo se deben establecer los parámetros de gobierno de datos y las mitigaciones requeridas que minimicen la probabilidad de estos riesgos. Esto es posible a través de acuerdos de nivel de servicio con proveedores de servicios para la identificación de fallas, reparación y recalibración de dispositivos a intervalos regulares; estándares aceptables para la calidad de los datos; requisitos de disponibilidad y tiempo de actividad; y coherencia de los datos a lo largo del tiempo.

#### Responsabilidad derivada de la calidad de los datos

Las instituciones productoras o generadoras de la información deben ser transparentes sobre cualquier problema potencial de calidad en los acuerdos de licencia o intercambio si los datos se pondrán a disposición de otros como datos abiertos o como datos compartidos.

Las instituciones deben informar a los usuarios que pueden existir situaciones de datos incompletos o cuestiones de calidad, con lo cual se salva la responsabilidad que pueda surgir por temas de calidad de datos. Sin embargo, en el marco de la gestión de la información sectorial, las instituciones adscritas deben implementar proceso de gestión de la calidad de los datos, y los lineamientos de gobierno de datos.

En etapas avanzadas de la gestión de información, en las cuales se han definido las responsabilidades de las instituciones en cuanto a la calidad de la información, se deben declarar los reportes de calidad en los metadatos.

### 3.2 Integración de datos

La integración de datos es el proceso de combinar datos de diferentes fuentes en una sola vista unificada. Con este fin, se deben definir y dar a conocer los requisitos para integrar datos de diferentes fuentes.

La integración comienza con el proceso de ingestión e incluye pasos como la limpieza, el mapeo y la transformación. En última instancia, la integración de datos permite el análisis de modelos federados de información para respaldar las herramientas de análisis que pueden producir inteligencia de procesos, eficaz y procesable.

Si bien, la integración es el proceso de tomar datos de varias fuentes dispares y hacerlos utilizables, en la medida en que aumenta el número de fuentes, la necesidad de una integración de datos eficaz se vuelve más importante.

Hay varios componentes que se deben tener en cuenta en la integración de datos:

#### Migración de datos

La **migración de datos** es el proceso de mover datos entre ubicaciones, formatos o aplicaciones. A menudo se debe a la introducción de un nuevo sistema o ubicación para los datos.

Una causa común en la actualidad es el cambio del almacenamiento y las aplicaciones locales al almacenamiento y las aplicaciones basados en la nube. Esto también es relevante, cuando los datos se transfieren a, y desde proveedores de servicios, ya sea para un proyecto (duración corta) o para un contrato de servicios o para subcontratar las operaciones y / o el mantenimiento de activos de información (duración más larga).

#### Integración de aplicaciones

La **integración de aplicaciones** busca lograr la interoperabilidad entre diferentes sistemas o aplicaciones.

Específicamente, requiere abordar problemas relacionados con la estructura organizativa de una institución y la implementación de acuerdos con otras instituciones. Algunos factores clave a considerar son:

- Interoperabilidad: administrar los diferentes sistemas operativos, incluidos los formatos de datos para que puedan conectarse;
- Integración: creación de un proceso estándar para gestionar el flujo de datos entre aplicaciones y sistemas para garantizar la coherencia; y
- Robustez, estabilidad, escalabilidad: independientemente de las soluciones implementadas, debe poder adaptarse a los cambios dentro del entorno institucional. Las soluciones típicas también incluyen middleware para ayudar con la centralización y estandarización de la gestión de datos.

### Agregación de datos

La federación de datos normalmente crea una base de datos virtual que no almacena los datos de origen, pero contiene información sobre dónde están los datos reales. Independientemente de cómo y dónde se almacenen los datos, deben presentarse como un conjunto de datos integrado. A menudo, esto implica transformación, limpieza y, si es necesario, enriquecimiento de datos.

El almacenamiento de datos agrega datos estructurados de una o varias fuentes con el fin de comparar y analizar los datos para lograr una mayor eficiencia en los procesos. Es eficaz para comprender mejor el rendimiento general de las aplicaciones y los activos asociados de información asociados, porque permite que una amplia gama de datos esté disponible para su análisis.

## 3.3 Seguridad de los datos

La protección de la confidencialidad y la integridad de los datos, al tiempo que se mantiene la disponibilidad y accesibilidad de los sistemas subyacentes, requiere una evaluación y gestión adecuadas de los riesgos de ciberseguridad.

Los riesgos de ciberseguridad son eventos que podrían dar lugar a acceso, uso, divulgación, interrupción, modificación o destrucción no autorizados de la información, la tecnología de la información y / o la tecnología operativa.

Los riesgos de seguridad cibernética deben considerarse como parte del entorno de riesgo de las instituciones y por tanto, debe alinearse con la estrategia y la práctica de gestión de riesgos de las entidades.

La toma de decisiones a lo largo del ciclo de vida de los datos, debe estar guiada por la gestión de riesgos para identificar mitigaciones y evitar riesgos que están fuera de la tolerancia al riesgo de la entidad y gestionar el riesgo de ciberseguridad, de acuerdo con las políticas y los Sistemas de Gestión de la Seguridad de Información, y las normas asociadas.

### Implementación de la privacidad por diseño

La privacidad por diseño y las evaluaciones del impacto de la privacidad pueden ayudar a garantizar la privacidad y la innovación y proporcionar una base sólida para que los datos se utilicen de forma anónima.

Mapear los flujos de datos y los activos de información, quién los posee y cómo se manejan en diferentes etapas del ciclo de vida del activo, puede ayudar a identificar cualquier riesgo de privacidad inherente al proyecto y a implementar la privacidad por diseño.

Es importante supervisar la creación, el uso y el acceso a los datos para garantizar un uso adecuado y seguro.



@upmeoficial



upme(oficial)



[www.upme.gov.co](http://www.upme.gov.co)