



La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Alejandra Bedoya Obando

**Trabajo de investigación presentado para optar al título
Magíster en Ciencia de la Información**

Asesores

Camilo García Morales

, Magíster en Gestión de Ciencia, Tecnología e Innovación

Andrés Sáenz Giraldo

**Magíster en Gestión Documental, Transparencia
y Acceso a la Información**

**Universidad de Antioquia
Escuela Interamericana de Bibliotecología
Maestría en Ciencia de la Información (V cohorte)
Medellín
2024**

Cita

(Bedoya Obando, 2023)

Referencia

Bedoya Obando, A. (2023). *La gestión de datos en Colombia: un estudio desde la perspectiva archivística* [Trabajo de investigación]. Universidad de Antioquia, Medellín.

**Estilo APA 7.ª edición
en inglés (octubre de
2019)**



Maestría en Ciencia de la Información, V Cohorte

Asesor: Camilo García Morales

Coasesor: Andrés Sáenz Giraldo



Biblioteca Carlos Gaviria Díaz

Repositorio Institucional

<http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes

Decano/director/a: Dorys Liliana Henao Henao

Jefe Departamento de Formación Académica: Camilo García Morales

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Contenido

Introducción	5
1. Planteamiento del problema	6
2. Justificación	8
3. Objetivos	10
4. Marco teórico	10
4.1 El ciclo de vida: datos y documentos.....	10
4.2 La necesidad de gestionar la información: datos y documentos	14
4.3 La archivística y los datos.....	17
5. Revisión bibliográfica	20
6. Metodología	26
6.1 Recolección de información.....	27
6.2 Análisis y presentación de información	29
7. Resultados	30
7.1 Características de la gestión de datos y la gestión documental.....	31
7.1.1 La gestión de datos.....	32
7.1.2 La gestión documental	46
7.2 Relaciones entre la gestión de datos y la gestión documental	57
7.2.1 Gestión de datos y gestión documental: definiciones	57
7.2.2 Tipos de datos y documentos	59
7.2.3 Gestión de datos y gestión documental: Principios	61
7.2.4 Gestión de datos y gestión documental: Procesos	66
7.3 Estrategias para la gestión archivística de datos en Colombia.....	72
7.3.1 Generación de datos	74
7.3.2 Valoración de datos.....	77
7.3.3 Preservación de datos.....	80
8. Conclusiones	83
9. Referencias bibliográficas	85

Lista de tablas

Tabla 1: Cantidad de bibliografía.....	28
Tabla 2: Ecuaciones para la búsqueda de bibliografía	29
Tabla 3: Criterios de inclusión y/o exclusión.....	29
Tabla 4: Tipos de datos y documentos	59
Tabla 5: Relación entre los principios.....	62
Tabla 6: Principios en la gestión de datos y la gestión documental.....	64

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Tabla 7: Gestión de datos y gestión documental: procesos estratégicos.....	67
Tabla 8: Gestión de datos y gestión documental: procesos técnicos.....	69

Lista de figuras

Figura 1 Principios de la gestión de datos.....	34
Figura 2 Procesos de la gestión de datos.....	38
Figura 3 Principios de la gestión documental.....	48
Figura 4 Procesos gestión documental.....	52
Figura 5 Gestión de datos y gestión documental: definiciones.....	58
Figura 6 Estrategias de la gestión archivística de datos en Colombia.....	74

Siglas y acrónimos

AGN	Archivo General de la Nación
Conpes	Consejo Nacional de Política Económica y Social
EIB	Escuela Interamericana de Bibliotecología
DPN	Departamento Nacional de Planeación
Icontec	Instituto Colombiano de Normas Técnicas y Certificación
ISO	International Organization for Standardization
NTC	Norma Técnica Colombiana
SGDEA	Sistema de Gestión de Documentos Electrónicos de Archivo
TIC	Tecnologías de la Información y las Comunicaciones
TRD	Tablas de Retención Documental
SIC	Sistema Integrado de Conservación

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Introducción

Esta investigación se realizó en el marco de la V cohorte de la Maestría en Ciencia de la Información, ofrecida por la Escuela Interamericana de Bibliotecología de la Universidad de Antioquia, sede Medellín; en coherencia, este ejercicio de escritura da cuenta de ello. La investigación surgió en vista de una problemática actual que refiere a la generación y producción masiva de datos, y a la necesidad de contar con estrategias para gestionarlos. En este sentido, se planteó comprender la gestión de datos desde la perspectiva archivística en Colombia para eso, se identificaron las características de la gestión de datos y la gestión documental, se analizaron las relaciones entre estos conceptos y, finalmente, se enunciaron algunas estrategias para la gestión archivística de datos en Colombia. En su desarrollo, se elaboró un análisis teórico acerca del ciclo de vida de los datos y los documentos; además, se abordó la necesidad de gestionar la información, la archivística y los datos, y para complementar los referentes teóricos, se desarrolló un apartado contextual donde se evidencian unos avances académicos y publicaciones, enfocados en la gestión de datos y cómo estos han sido tratados desde la disciplina archivística.

En la metodología se definió el enfoque cualitativo, ya que en el proceso se recolectaron datos descriptivos; de allí el tipo de investigación. Respecto del método, se ejecutó la investigación documental, en razón a esta facilita el rastreo bibliográfico en la obtención de la información necesaria, lo que favorece la materialización de cada objetivo planteado.

Fue así como se lograron evidenciar las características de la gestión de datos y la gestión documental, sus definiciones, tipos de datos y documentos, principios y procesos —tanto estratégicos como técnicos—, información fundamental para realizar la comparación, donde se determina una complementariedad entre la gestión documental y la gestión de datos. Finalmente, se plantearon tres estrategias para la gestión archivística de los datos en Colombia y, para cada una de estas, se definieron y desarrollaron unos componentes con sus respectivas actividades.

Por lo expuesto hasta aquí, es dable concluir que a lo largo de esta indagación se logró plantear una alternativa a la problemática en torno a la producción masiva de datos, la cual puede

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

ser controlada con las estrategias para la gestión archivística de los datos, esbozadas en el desarrollo de esta investigación, ya que con estas se busca administrar los datos durante todo su ciclo de vida.

1. Planteamiento del problema

En la actualidad, se están generando gran cantidad de datos y diversas manifestaciones documentales —como textos, imágenes y videos—, facilitados por las tecnologías y su interacción con las personas; de ahí que «nuestra capacidad de generación de datos nunca ha sido tan poderosa y enorme desde la invención de la tecnología de la información a principios del siglo XIX» (Wu *et al.*, 2013, p. 2), tanto que el «volumen de datos generado en el mundo en 2018 —fue— calculado en 33 zettabytes y se estima que será de 175 zettabytes en 2025» (Amutio, 2021, p. 101).

Si esta producción de datos «no dispone de un adecuado sistema de gestión de todo su ciclo de vida, se corre el peligro de una mala conservación de grandes masas de datos y documentos cada vez más inmanejables» (García, 2013, p. 3). Como alternativa a esta situación, en Colombia se cuenta con una serie de normas que regulan la gestión documental, como son: Ley 594 de 2000, Decreto 1080 de 2015, «por medio del cual se reglamenta el Título V de la Ley 594 de 2000, parcialmente los artículos 58 y 59 de la Ley 1437 de 2011, y se dictan otras disposiciones en materia de Gestión Documental para todas las Entidades del Estado». Además, se han publicado una serie de lineamientos, manuales y programas, entre los cuales se encuentran el Programa de gestión documental, la Guía para la anonimización de datos, el Modelo de requisitos para un sistema de gestión de documentos electrónicos de archivo, el Sistema integrado de conservación, entre otros recursos facilitados por el Archivo General de la Nación (en adelante, AGN), cuyo propósito es orientar a las entidades, de manera que se garantice una correcta gestión documental.

Cabe resaltar que dentro de estos recursos, y en la normatividad en materia de gestión documental, se han contemplado los datos de manera muy específica; por ejemplo, en la Guía de implementación de un sistema de gestión de documentos electrónicos de archivo se menciona que «si la información electrónica es un “documento”, un “documento de archivo”, un “dato” o “contenido”, la concepción de la información es la misma, es “electrónica” y debe ser gestionada por al menos un componente tecnológico» (AGN, 2020).

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

En esta guía se contempla la *captura* como uno de los procesos de un sistema de gestión de documentos electrónicos, donde se «requiere llevar a cabo actividades de planeación y alistamiento de información contemplando los principios archivísticos, y puede abarcar el uso de tecnologías para el reconocimiento y clasificación de documentos, herramientas de extracción de datos automáticamente y almacenamiento de metadatos» (AGN, 2020). Lo anterior, para cumplir el fin de los archivos que consiste en garantizar el «servicio de la gestión administrativa, la información, la investigación y la cultura» (Congreso de la República de Colombia, 2000).

Ahora bien, para el caso de la gestión de datos en Colombia, se cuenta con la Política nacional de explotación de datos, publicada en el año 2018 por el Departamento Nacional de Planeación (en adelante, DNP), la Guía técnica de información sobre gobierno del dato (2019) y la Guía para el uso y aprovechamiento de datos abiertos (2019), publicadas por el Ministerio de las TIC ese mismo año.

Para ejecutar la gestión de datos, se requiere «la elaboración de procedimientos metodológicos y tecnológicos para su preservación y el desarrollo de normas y criterios para su uso, almacenamiento y disponibilidad» (Martínez, 2021), así como «asegurar que el marco legal y las políticas de protección de datos, seguridad, interoperabilidad, y propiedad intelectual sean compatibles con los datos» (Amutio, 2021, p. 109). El desafío de gestionar los datos implica «saber el límite de usabilidad de los datos y poder responder preguntas como ¿quién es el dueño de los datos?, ¿se puede garantizar la integridad de los datos?, ¿cuál es límite en el uso de datos por parte de las organizaciones?, ¿qué datos son privados?» (Arriagada, 2020, p. 2), además de «lograr cierta claridad en las definiciones de los puntos en común y las diferencias entre "datos" y "documentos"» (Cunningham, 2021, p. 32).

La posibilidad, desde el campo archivístico, para abordar los datos de manera que se asegure el desarrollo de una articulación y complemento entre las políticas de gestión de datos y de documentos, implica la estandarización de metodologías o guías que permitan a las entidades una gestión de información integral, de modo que no solo se contemplen los documentos que se tienen en formato papel y electrónico, sino también las demás representaciones documentales, audiovisuales, correos electrónicos, y los datos como información producida, recibida o capturada por las instituciones, en virtud de sus funciones. Lo anterior alerta a la disciplina archivística a comprender el funcionamiento para la gestión de los datos, desde sus diferentes teorías, procesos, procedimientos e instrumentos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

La gestión de datos, desde la perspectiva archivística en Colombia, es un tema que aún no ha sido desarrollado, más allá de unas aproximaciones concretas; por esta razón, es necesario que se contemplen los datos y los documentos como parte de los recursos de información que producen, gestionan y conservan las entidades, pues el hecho de manejar la información de manera fraccionada, dificulta el proceso de consulta, promueve escenarios de riesgo — como la pérdida de la información y la manipulación inadecuada—, obstaculiza el acceso y hace que las respuestas a usuarios (personal interno de la entidad, usuarios externos, sociedad en general, entre otros) sean menos diligentes, haciendo que el derecho al acceso a la información, con los requisitos de agilidad y veracidad de esta, no se cumpla a cabalidad. Aquí, precisase origina la pregunta de investigación: ¿cómo se puede ser comprendida la gestión de datos en Colombia desde la perspectiva archivística?

2. Justificación

La investigación se desarrolla en el marco de la maestría en Ciencia de la Información, con énfasis en gestión de los datos, la información y el conocimiento en el entorno digital, uno de los programas de posgrado brindado por la Escuela Interamericana de Bibliotecología (EIB), que busca

formar profesionales idóneos para asumir los desafíos que les demanda la sociedad de la información actual, así como los cambios sociales generados por las prácticas actuales en cuanto al uso de la información y sus nuevas formas de representación, mediadas por las tecnologías y la información digital. (Universidad de Antioquia, 2021, p. 1).

Cabe resaltar que la Ciencia de la Información «tiene sus antecedentes en la aplicación de la tecnología para el análisis y procesamiento de los datos» (Martínez, 2021, p. 32), lo que abre paso al tema de investigación que nos ocupa.

La gestión de información es una actividad que se realiza de manera transversal en las entidades, que goza de gran relevancia a la hora de la toma de decisiones; valga advertir, requiere de control, pues la cantidad de información que se produce y gestiona en cada institución tiende a crecer exponencialmente, de allí la importancia y la necesidad de gestionar la información de una manera integral; es decir, no solo contemplando los documentos que se tienen en formato papel y electrónico, sino también las demás representaciones documentales, audiovisuales, correos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

electrónicos; adicionalmente, datos como: información producida, recibida o capturada por las instituciones o personas naturales, todo ello, en virtud de sus funciones. En este sentido,

los especialistas y profesionales en gestión de documentos debemos tener en cuenta que muchas de las cosas que sabemos, o que hemos aprendido a hacer, se pueden poner en valor y aplicar en nuevos entornos; algunas cosas debemos adaptarlas y transformarlas para que sigan siendo útiles y otras tenemos que aprenderlas de nuevo, consolidando nuevas buenas prácticas que puedan ser aplicadas. (Bustelo, 2021, p. 77).

Es pertinente que los archivistas o profesionales de la información hagan parte del grupo de personas capacitadas que posibiliten la gestión de información integral en cada entidad, dando cabida a la gestión de los datos que, al igual que los documentos, pueden ser tratados, manejados y administrados desde la perspectiva archivística.

Ahora bien, la gestión de los datos ha sido estudiada desde la ciencia de datos, y se ha tratado desde las áreas de las TIC, responsables del tratamiento de estos. Esta investigación, por ende, pretende analizar la gestión de los datos desde la perspectiva archivística, una disciplina perteneciente a las ciencias de la información. Este análisis es realizado teniendo en cuenta que el concepto *dato*, de acuerdo con las búsquedas bibliográficas realizadas en esta investigación, ha sido tratado de manera insuficiente dentro de la normatividad, manuales, producción académica y guías en materia de gestión documental en Colombia.

Por otra parte, del concepto *gestión archivística de datos* hay que decir que se ha venido tratando recientemente; por ejemplo, en los últimos años fue reglamentado en Cataluña (España) mediante el Decreto 76 del año 2020, donde se menciona que la finalidad de la gestión archivística de datos es «asegurar la creación, la captura y la gestión de una información auténtica, fiable, usable e íntegra que permita el desarrollo de las actividades y competencias» (Departamento de Políticas Digitales y Administración Pública, 2020, p. 15).

Partiendo de la premisa de que «cuando en una entidad se habla de gestión de datos y gestión documental, en realidad se está hablando de lo mismo: gestión de la información» (Serra, 2021a), se da paso a una nueva «cuestión de si deberíamos preocuparnos o no por la gestión de datos o si ese desafío en particular es asunto de otros, no el nuestro» (Cunningham, 2021), haciendo referencia a la disciplina archivística y los profesionales de esta; sin embargo, “claramente tiene

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

nuevas implicaciones para nuestra profesión que debemos considerar, implicaciones que [...] deberían hacernos repensar lo que parece haber sido un punto ciego profesional durante demasiado tiempo, remata Cunningham (2021, p. 32).

De acuerdo con lo expuesto inmediatamente anterior, el tema de investigación sobre gestión de datos es pertinente para la disciplina archivística, desde su área profesional y académica; además, teniendo en cuenta las dinámicas de generación de información que se tienen actualmente en la sociedad y en las entidades, esta facilita la solución de inconvenientes que se han presentado, o se presentarán en las instituciones, donde la gestión de información es un proceso de ejecución diaria y necesaria para su funcionamiento en todas las áreas que la conforman.

3. Objetivos

3.1 General

Comprender la gestión de datos desde la perspectiva archivística en Colombia para su tratamiento.

3.2 Específicos

- Identificar las características de la gestión de datos y la gestión documental desglosando cada concepto con el fin de obtener las bases necesarias para la comprensión de cada uno de ellos.
- Analizar las relaciones entre la gestión de datos y la gestión documental desde la perspectiva archivística.
- Proponer estrategias para la gestión archivística de datos en Colombia.

4. Marco teórico

4.1 El ciclo de vida: datos y documentos

En diferentes áreas del conocimiento, como las matemáticas, la estadística, la informática y las ciencias de la información, ha sido usado el término dato, lo que muestra un uso amplio de este concepto. Cabe resaltar que este ha ido adaptándose y evolucionando a medida que ocurren los avances tecnológicos. Dentro de este apartado se hace mención de las diferentes definiciones que sobre las acepciones dato y documento hay, al igual que el ciclo de vida por el que estos atraviesan.

En primer lugar, el concepto *dato*, que comenzó siendo usado en el área de las matemáticas, es definido como «un hecho dado como base para el cálculo en problemas matemáticos» («Online

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

etymology dictionary», 2022). Posteriormente pasó a ser incluido dentro del concepto de información, ya que para el año 1727, de acuerdo con lo expuesto por Iramain (2000), se hizo «uso por primera vez de la palabra *information* para referirse, sin más matices ni aclaraciones, a datos estáticos: no ya al acto mismo de entender, como se venía utilizando hasta el momento, sino al producto de la aprehensión intelectual» (p. 11), teniendo en cuenta lo afirmado por Yeo en el año 2018, cuando aseguró que «los datos son transformables en conocimiento, que luego se puede compartir y convertir en información» (p. 11).

En relación con los conceptos dato e información, los «data has been called the “raw material of information” and information has been called “data in context» (Global Data Management Community, 2017) [los datos han sido llamados la “materia prima de la información” y la información ha sido llamada “datos en contexto”] (Global Data Management Community, 2017, p. 51). Ahora bien, “inrelation to information technology, data is also understood as information that has been stored in digital form” [en relación con la tecnología de la información, también se entiende por datos la información que ha sido almacenada en forma digital].

El dato también ha sido usado en la estadística; se dice que dicha área «es la primera ciencia que considera los datos como su materia prima, pero las nuevas necesidades y, en particular, las nuevas características de los datos (en volumen y tipología) hacen que las disciplinas que integran lo que se conoce como "minería de datos" sean numerosas y heterogéneas» (Hernández y Ramírez, 2004), entendiendo la minería de datos como el proceso de extracción de conocimiento a partir de los datos, el cual es ejecutado por un grupo interdisciplinario.

Este concepto ha sido tratado, igualmente, en el área de la informática como aquel «contenido que presentan las computadoras a los usuarios» (Yeo, 2018, p. 111), ya sea en «conjuntos de datos o bases de datos relacionales; se incluye el material audiovisual y los datos alfanuméricos [...] extraídos o derivados de contenidos de registros organizacionales» (Yeo, 2018, pp. 111-112), siendo las bases de datos relacionales aquellas que poseen «un esquema asociado, es decir, los datos deben seguir una estructura y son, por tanto, estructurados» (Hernández y Ramírez, 2004).

En informática también se encuentra el concepto *datos digitales*, interpretados mediante codificación binaria, los cuales «pueden generarse de dos maneras: por la interacción de personas con sistemas, herramientas y servicios digitales o automáticamente por programas de *software* y dispositivos de *hardware* que los capturan» (DPN, 2018).

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

En el área de la archivística, una disciplina de las ciencias de la información, el dato comenzó siendo usado en las herramientas de descripción y consulta; según Martínez (2021), «los profesionales de los archivos, si en algo hemos destacado, es en crear y desarrollar normas de datos, aunque las denominemos estándares de descripción» (p. 88); es por ello que en estas herramientas propias de la disciplina, más concretamente de la gestión documental, se extraen y disponen datos para facilitar la identificación y consulta de las unidades documentales.

La visión de dato planteada anteriormente se ha ido transformando, de una parte, de acuerdo con los avances que se han dado en la disciplina archivística; de otra, de acuerdo con varios cuestionamientos sobre la «necesidad de repensar el concepto de documento al que se añaden hoy otras manifestaciones documentales como páginas webs, plataformas, bases de datos, etc.» (Heredia, 2018, p. 8); además, «se nos avisa que nos quedaremos atrás si gestionamos solo documentos digitales o electrónicos, porque la información a partir de los datos cada vez es mayor y con más valor», por cuanto la tendencia es «a la desmaterialización de las series convirtiéndolas en datos integrantes de grandes bases de datos», concluye Heredia (2018, p. 15).

En esta misma línea, dentro de la disciplina archivística se pueden encontrar otros referentes del concepto, como son: «representación de todo y cualquier elemento de contenido cognitivo, susceptible de ser comunicado, procesado e interpretado de forma manual o automática» (Mastropiero, 2008, p.) o como aquella «unidad mínima, que sólo se convierte en “información” cuando los datos transmiten un mensaje o tienen un sentido» (Bustelo, 2021, p. 68).

Para el caso de Colombia, los datos han sido definidos como «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» (Ministerio de Tecnologías de la Información y las Comunicaciones, 2019).

Una definición más es brindada por el DPN (2018) que menciona los datos como un recurso de infraestructura; es, dice,

un insumo para la actividad investigativa, académica y productiva, para la toma de decisiones públicas y para la garantía de derechos [...] distintos a otros activos económicos en la medida en que su valor se materializa con el procesamiento y obtención de la información y el conocimiento intrínseco que se obtiene al combinarlos entre sí. (p. 32).

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Ahora bien, dentro de la Política nacional de explotación de datos, se mencionan los tipos de datos

estructurados, organizados conforme a un modelo o esquema; no estructurados, en los que su organización y presentación no está guiada por ningún modelo o esquema y semiestructurado, su organización y presentación tiene una estructura básica (etiquetas o marcadores), pero no tiene establecida una definición de relaciones en su contenido. (DNP, 2018).

Conexo a lo dicho, también se aluden unas tipologías de datos de acuerdo con su criterio orgánico, es decir, el sujeto que genera el dato, los cuales pueden ser públicos o privados (DNP, 2018).

Por otro lado, en Colombia, los documentos son considerados «información registrada, cualquiera que sea su forma o el medio utilizado» (AGN, 2006); por lo tanto, los documentos pueden ser análogos, ofimáticos, electrónicos o digitales. Así mismo, un documento es

un tipo de soporte que utiliza la fuente para registrar la información, para que pueda ser mejor difundida (de forma escrita, sonora, simbólica, etc.). A diferencia de la fuente, el documento puede ser procesado (catalogado, indizado, etc.) y preservado según el tipo de material (papel, plástico, etc.). (Pérez, 2009, p. 7).

Cruz Mundet (2011) aporta a la discusión manifestando que los elementos que caracterizan los documentos son «el soporte que le confiere corporeidad física, y puede ser desde una tablilla de barro hasta un disco óptico [...] es decir, la noticia que transmite y el registro, o sea, la fijación de la información en el soporte, ya sea mediante tinta, impulsos electromagnéticos u otro cualquiera» (p. 15).

Cabe anotar que los documentos tienen un ciclo de vida (considerado como una serie de etapas), tal como lo afirma Cruz Mundet (2011): «tiene una vida similar a la de un organismo biológico, el cual nace (fase de creación), vive (fase de mantenimiento y uso) y muere (fase de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

expurgo)» (p. 7). Por lo tanto, el ciclo vital del documento de archivo se define como aquellas «etapas sucesivas por las que atraviesan los documentos desde su producción o recepción, hasta su disposición final»; así mismo, el ciclo de vida del documento electrónico se define como aquel «formado por las distintas fases por las que atraviesa un expediente desde su creación hasta su disposición final» (AGN, 2012). Es en este sentido que «en el entorno de los archivos electrónicos el ciclo de vida tiene tres etapas básicas: diseño, creación y mantenimiento (incluye preservación y uso)» (Zea y Pulido, 2016, p. 13).

Cabe indicar que, del mismo modo, los datos tienen un ciclo de vida «based on the product lifecycle [...] includes processes that create or obtain data, those that move, transform, and store it and enable it to be maintained and shared, and those that use or apply it, as well as those that dispose of it» [basado en el ciclo de vida del producto [...]; incluye los procesos que crean u obtienen datos, los que los mueven, transforman y almacenan y permiten para ser mantenido y compartido, y los que lo utilizan o aplican, así como los que disponen de eso] (Global Data Management Community, 2017, p. 58), desde su creación hasta su eliminación.

Está definido como una serie de «procesos que van desde la generación de información hasta que esta se convierte en conocimiento capaz de generar una nueva información, ya sea a nivel individual de las personas como a un nivel social» (Caraballo y Ramírez, 2007, p. 32), este ciclo de vida de los datos «se compone de cuatro etapas fundamentales: modelado del dato, captura del valor, almacenamiento y visualización» (Domingo *et al.*, 2007, p. 586), que abarcan la recopilación, preservación y uso de los datos.

4.2 La necesidad de gestionar la información: datos y documentos

La necesidad de gestionar la información surge como consecuencia de la producción de datos y documentos en un mundo que se transforma continuamente, en coexistencia con lo digital, con la gran acogida que tienen las herramientas informáticas y con la facilidad de compartir y consumir información. Esta gestión busca garantizar la administración y el uso de los recursos como datos y documentos para la toma de decisiones en las organizaciones o entidades. En primer lugar, qué se entiende por *información*, según Capurro (2014), se trata de la «acción de dar forma a algo material, así como el de comunicar conocimiento a una persona» (p. 6); de ahí que se le otorgue la capacidad de transmitir mensajes; adicionalmente, se indica que esta tiene que ver con su capacidad de percepción y conocimiento del mundo. Agrega Capurro que, de una parte, esta noción consta de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

dos categorías: «aquello que puede ser comprendido y aquello que genera información»; de otra, acota que la información circula en torno a tres paradigmas: el físico, el cognitivo y el social, modelo que la convierte en algo tangible, plena de conocimiento, resultante de una construcción social.

En segundo lugar, la *gestión de información* se plantea como aquella alternativa que se implementa para dar orden y uso a los documentos y datos generados en las instituciones. Ahora bien,

si se intenta desprender el sistema de gestión de información del sistema de gestión documental, no se aprovechará toda la información por lo que el manejo de información se verá desprovisto del respaldo de las fuentes, lo que limitará la combinación de informaciones y hará que la organización carezca de la fortaleza que implica el respaldo de sus operaciones y actuaciones. (Ponjuan, 2005, p. 4).

A raíz de lo planteado anteriormente, se debe tener en cuenta que la información también se encuentra materializada en documentos de archivo definidos como

una unidad de información registrada cuya amplitud abarca el documento en soporte papel, el electrónico, el administrativo, el de conservación permanente y es testimonio y prueba, siempre relacionado tanto con un agente con competencias atribuidas como con una función/actividad derivada de esas competencias. (Herrera, 2019, p. 67).

De esta información, se tiene, entonces, la articulación con la gestión documental, ya que es un «conjunto de normas, técnicas y conocimientos aplicados al tratamiento de los documentos desde su diseño hasta su conservación permanente» (Cruz Mundet, 2011, p. 18).

Sobre la gestión de datos, desde la perspectiva archivística, se vislumbra un «panorama prometedor para la gestión documental, puesto que si convenimos que todos los documentos están compuestos de datos, y que cualquier agregado significativo de datos es en el fondo un documento, la gestión de datos pasa a ser, en esencia, gestión de documentos» (Serra, 2021c, p. 47); de aquí, que «es el momento [...] de adoptar el análisis de los datos y la visualización de dichos datos como un proceso propio de la gestión documental», concluye.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Sumando a lo expuesto por Serra, la Global Data Management Community (2017) se aproxima a la definición de la siguiente manera: «is the development, execution, and supervision of plans, policies, programs, and practices that deliver, control, protect, and enhance the value of data and information assets throughout their lifecycles» [es el desarrollo, ejecución y supervisión de planes, políticas, programas y prácticas que entregan, controlan, protegen y mejoran el valor de los datos y los activos de información a lo largo de sus ciclos de vida] (p. 49).

En la dirección del ciclo de vida, atinente a los datos, agrega la misma asociación: «the specifics of the data lifecycle within a given organization can be quite complicated, because data not only has a lifecycle, it also has lineage (i.e., a pathway along which it moves from its point of origin to its point of usage, sometimes called the data chain)» [los detalles del ciclo de vida de los datos dentro de una organización determinada pueden ser bastante complicados, porque los datos no solo tienen un ciclo de vida, también tienen un linaje (es decir, un camino a lo largo del cual se mueven desde su punto de origen hasta su punto de uso, a veces llamado cadena de datos)] (p. 58).

Referido al uso, el Mintic (2018) asegura que la gestión de datos «utiliza el directorio de datos, el cual permite acceder, desde un único punto, a los distintos sitios Web y recursos del sector público que ofrecen información pública y de intercambio» (Ministerio de Tecnologías de la Información y las Comunicaciones, 2018); con ello se da cumplimiento a la finalidad, esto es, «asegurar, mantener y proveer instituciones de datos del sector público, unificando datos maestros y regularizando registros en los sistemas fuente», garantizando, así, el acceso a datos «organizados y estructurados por formatos y temas, entre otros criterios, como resultado de la aplicación del gobierno del dato». A este propósito, en Colombia,

las disparidades en los avances y capacidades para la gestión de datos de las entidades son la principal limitación, que se explica porque no todas tienen el mismo avance en la implementación de sus sistemas de información, ni la misma capacitación para gestionar información. (Departamento Nacional de Planeación, 2018).

El DNP, en su Política nacional para la explotación de datos (2018), realizó una serie de recomendaciones al Departamento Administrativo de la Función Pública, dentro de las cuales se encuentra la necesidad de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

proponer un modelo de gestión de datos que se articule con las políticas de gestión y desempeño institucional en el marco del modelo integrado de planeación y gestión de las entidades públicas, incorporar temáticas relacionadas con la explotación de datos en los planes de capacitación de las entidades públicas para fortalecer el capital humano del ecosistema de datos. (p. 94).

En palabras del DNP, las recomendaciones aludidas se realizan teniendo en cuenta que «la adopción de una política de gestión de datos por parte de las entidades es un mecanismo que podría solventar la incertidumbre y aportar estandarización en el tratamiento de las situaciones» (p. 55), así como la «multiplicidad de registros administrativos y reportes de información, debilidades archivísticas, desaprovechamiento de datos actualmente disponibles y barreras de acceso a los datos públicos» (p. 53).

4.3 La archivística y los datos

Adentrándonos en el terreno particular de la archivística, es definida por Cruz Mundet (2011) como la «ciencia de los archivos [...] integrada por un conjunto de conocimientos y de métodos para el tratamiento de los documentos y de los archivos» (p. 16) que, junto a la ciencia de datos, se identifican como «disciplinas que analizan, entre otros, la producción y diversidad documental, la explosión de la información, el análisis masivo de los datos» (Martínez, 2021).

En Colombia la archivística es asumida como una «disciplina que trata los aspectos teóricos, prácticos y técnicos de los archivos» (AGN, 2006). Por su parte, Fuster (1999) deja claros los principios fundamentales que marcan la diferencia entre la archivística y las demás ciencias de la documentación, a saber: «el carácter seriado de los documentos, el sentido del archivo como conjunto orgánico de documentos y nunca como colección documental, y el principio de procedencia» (p. 105); agrega, además, que «la Archivística está considerada como una más de las llamadas Ciencias de la Información, adoptando algunas de las herramientas y técnicas de trabajo de las mismas, en especial la creación de bases de datos y la aplicación de la informática» (p. 120).

Adicionalmente, se habla de la *archivística integrada*, propuesta por los canadienses Couture *et al.* (1988; citados por Betancur y Jaramillo, 2021) en la década del ochenta, quienes hacen referencia a una «disciplina cuya razón de ser está en el corazón de la gestión de la información, recurso vital de la organización» así las cosas, se «supone que la gestión de documentos debe incorporarse a los fundamentos tradicionales del área; consideran que no son dos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

disciplinas distintas, sino que su fusión reforzará la tradición archivística basada en los principios» (Varela y Saraiva, 2015, p. 113). Además, busca garantizar «el tratamiento de los documentos desde sus orígenes en el proceso administrativo hasta su conservación definitiva, lo que implicaba aglutinar el ciclo de vida» (Aguilera, 2012), es decir, la integración de las fases de archivo de gestión, central e histórico, como parte de un todo.

Conviene añadir que, con la llegada de las tecnologías de la información, la archivística ha experimentado grandes cambios; tanto es así que algunos autores identifican una archivística computacional, teniendo en cuenta que «computational methods and tools can be applied to perform archival functions» [se pueden aplicar métodos y herramientas computacionales para realizar funciones de archivo] (Marciano *et al.*, 2018). En este sentido, continúa el autor,

archival science without computer science lacks knowledge necessary to understand and effectively administer digital records; at the same time, computer science without archival science lacks knowledge needed to generate enduring and trustworthy memory and evidence – hence, the need for a combination of the two disciplines into the new transdiscipline [la archivística sin informática carece de los conocimientos necesarios para comprender y administrar eficazmente los documentos digitales; al mismo tiempo, la informática sin archivística carece del conocimiento necesario para generar memoria y evidencia duraderas y confiables; de ahí la necesidad de una combinación de las dos disciplinas en la nueva transdisciplina]. (p. 3).

Como lo aclara García (2017), «la irrupción de la informática permitió digitalizar procesos y, por tanto, gestionar datos en aplicaciones informáticas, concretamente en bases de datos relacionales, donde los datos están estructurados, o en aplicaciones que permitían explotar el contenido» (p. 4); es por esta razón, afirma García, que «la informática y la evolución de las aplicaciones nos han permitido tener en formato electrónico los datos y la información que generamos en nuestras actividades y aumentar la velocidad de procesamiento»(p. 5).

Dentro de la archivística emergen nuevos conceptos o paradigmas como la gestión documental datificada, en la cual

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

no se crea nada que previamente los gestores no tuvieran ya, sino que tan solo se sistematiza y normaliza algo que anteriormente ya existía de forma natural. Por lo tanto, la gestión datificada es la evolución natural de algo ya existente en todas las organizaciones. (Serra, 2021b, p. 522).

En efecto, acota Serra,

la puesta en marcha de un modelo de gestión orientado a datos tiene como consecuencia inmediata una aceleración en el ciclo de actividad de la organización: nuevas oportunidades, nuevos productos o servicios, mayor *feedback* sobre los efectos en el negocio, mayores posibilidades de mejora, y vuelta a empezar. —Lo anterior tiene como propósito fundamental— mejorar las decisiones que se toman en el curso de cada proceso y las condiciones en que se toman. En esta toma de decisiones solo importan los datos y su fiabilidad; si los datos se autentican, el documento puede acabar convirtiéndose en una mera formalidad. (p. 524).

Con todo lo dicho va quedando claro que la gestión de datos, vista desde la disciplina archivística, fue tratada por Jordi Serra, quien manifiesta:

Se tienen dos figuras trabajando con lo mismo, los gestores de documentos, que se encargan del manejo de objetos discretos, es decir, datos no estructurados, y los gestores de datos, enfocados en la información continua o estructurada; sin embargo, esos datos continuos ya son estructurables, incluso los no estructurados; por lo tanto, la gestión de datos no solo se enfoca en la gestión de datos, sino también de [*sic*] estructurar datos.

Por ello, advierte Serra, se establece que la relación entre gestión de datos y gestión documental se hace muy estrecha, pues «cuando en una entidad se habla de gestión de datos y gestión documental, en realidad se está hablando de lo mismo: gestión de la información».

Es importante resaltar que recientemente la gestión archivística de los datos fue reglamentada en España, particularmente en la Comunidad Autónoma de Cataluña, mediante el Decreto 76 de 2020, donde se define que con esta se busca «asegurar la creación, la captura y la gestión de una información auténtica, fiable, usable e íntegra» (Departamento de Políticas Digitales y Administración Pública, 2020); no obstante, «todavía queda el desarrollo de este decreto para ver cómo se va a concretar; pero este reconocimiento a nivel legislativo es todo un hito que esperamos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

que se replique en otras Comunidades Autónomas y, por qué no, a nivel estatal también» (Bustelo, 2021); por lo tanto, partiendo de la fuerte influencia de la teoría archivística española en Colombia, es factible proyectar que la gestión archivística de datos será también replicable en nuestro territorio, no perdiendo de vista que «cualquier sistema archivístico debe disponer de un plan o una estrategia que aborde la política de gestión de datos que se manejan para poder valorar los documentos que se han de conservar o los datos que se han de mantener» (López y Requejo, 2021).

De acuerdo con lo definido por la Comunidad Autónoma de Cataluña, «la gestión archivística de los datos y de los activos digitales tiene como finalidad asegurar la creación, la captura y la gestión de una información auténtica, fiable, usable e íntegra que permita el desarrollo de las actividades y competencias propias de los sujetos» (Departamento de Políticas Digitales y Administración Pública, 2020); en esta se desarrollan tres fases:

la creación de la información digital hasta la finalización de la tramitación o asunto que se gestiona, la conservación de la información, mientras los documentos o los datos conserven el valor probatorio de acciones, derechos y deberes de la Administración o de las personas y la conservación de la información digital que ha sido dictaminada de conservación permanente y la destrucción del resto de información de acuerdo con la normativa vigente en esta materia.

5. Revisión bibliográfica

En el curso de esta búsqueda, y para la elaboración de este apartado, se revisó la literatura disponible y relevante, mediante la cual se identificaron los avances relativos a la gestión de datos, desde la perspectiva archivística. Fue así como se realizaron búsquedas en bases de datos bibliográficas como Scielo, ScienceDirect, Scopus, Springerlink, Redalyc. Las mismas, se hicieron, en primer lugar, bajo las categorías de ciclo de vida de los datos y los documentos, gestión de información, archivística y datos; igualmente, se emprendieron otras con palabras clave afines a estas categorías, tales como: archivística en la era digital, gestión documental datificada, gobernanza de información, gestión de activos de información, gestión de información digital, gestión archivística de datos, gestión documental orientada a los datos y gestión de datos y archivistas. Es significativo, entonces, que, a partir de estas indagaciones, se haya realizado un análisis selectivo.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

En el año 2002, se publicó la investigación titulada *la Formación en Recuperación de Información: Análisis de los cursos y asignaturas en las escuelas de Biblioteconomía y Documentación de Norteamérica y España*, realizada por el profesor Salvador Oliván, de la Universidad de Zaragoza. El objetivo fue claro: buscaba

describir y analizar los cursos y asignaturas con contenido en recuperación de la información impartidos durante el curso 2001-2002 en las Escuelas de Biblioteconomía e Información en Norteamérica y en los Planes de Estudio de Biblioteconomía y Documentación de las Facultades y Escuelas de las universidades españolas. (p. 7).

Dentro de los hallazgos de esta indagación, se logró evidenciar que la evaluación de bases de datos y la capacidad de búsqueda en las mismas, con la finalidad de recuperar información, son cursos que se dictan en las escuelas de biblioteconomía y documentación en Norteamérica. Para el caso de España, esta también cuenta con varios cursos relacionados, verbigracia, el curso de Bases de Datos Documentales y Administración de Bases de Datos. Si bien la investigación no mostró la relación exacta entre la gestión de datos y la archivística, sí brindó una aproximación acerca de la inmersión de los datos o, en este caso, bases de datos dentro de la práctica archivística; lo anterior, teniendo en cuenta que dentro de la gestión de datos se encuentra la recolección, uso y, de manera general, la administración de los datos durante su ciclo de vida; adicionalmente, las asignaturas que se mencionan en esta investigación acercan a los archivistas a estas actividades, lo que nos indica que esta aproximación de los datos a la archivística, desde la perspectiva académica, no es aislado.

En el año 2017, la jefa de Publicación de Datos y Software en F1000, Rebecca Grant, quien se formó en la School of History and Archives, University College Dublin, publicó los resultados de la investigación *Recordkeeping and research data management: a review of perspectives* [*Gestión documental y gestión de datos de investigación: una revisión de las perspectivas*], enfocada, principalmente, en hallar la relación entre los datos de investigación y la gestión documental; así mismo, el papel de los archivistas dentro de las bibliotecas estadounidenses, donde se conservan los datos de investigación. La metodología implementada consistió en un rastreo de información bibliográfica, partiendo de una tesis de doctorado, *Professional records and research data* [*Profesionales de documentos y datos de investigación*], realizada en el año 2014 y nombrada como la primera parte del proceso.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Esta segunda etapa se enfocó en determinar la relación y el nivel en que los documentos y profesionales de la información se relacionan con los datos de investigación. Las conclusiones a las que se llegó en esta investigación fueron: en primer lugar, no se logró determinar si los profesionales de documentos son los más idóneos para gestionar los datos de investigación, en virtud a que «given the experience and expertise of records professionals, there is a gap in the literature regarding a potential role of records professionals as leaders in the development of research data management» [dada la experiencia y experticia de los profesionales de documentos, hay un vacío en la literatura con respecto al papel potencial de los profesionales de documentos como líderes en el desarrollo de la gestión de datos de investigación] (Grant, 2017, p. 18). A todo esto, mencionó que estos tienen cierta relevancia a la hora de gestionar los datos e, incluso, se citan como necesarios para ello; recordó, de igual modo, que el papel de los bibliotecarios y archiveros es valioso a la hora de prestar un servicio de datos de investigación.

Referido al contexto colombiano, en el año 2020 varios profesionales, realizaron la investigación denominada *Importancia del big data en un gestor documental para las entidades públicas de Colombia*, la cual se enfocó, principalmente, en «indagar cómo en las entidades públicas los sistemas de gestión de documentos electrónicos de archivo permiten identificar si el *big data* contribuye a satisfacer las necesidades de gestión de información» (Vargas *et al.*, 2020). La metodología implementada consistió en un estudio exploratorio cualitativo; se abordó una revisión sistemática de la información y, parte de los resultados, mostró la inexistencia de un proyecto de este tipo, es decir, donde se tuviera el *big data* en un gestor documental. De ahí que se haya concluido que

el primer paso para un correcto uso del *big data* en la gestión documental es la centralización e ingesta de información. La integración de datos y la fusión o implementación de estas herramientas son consideradas totalmente viables en Colombia, como una combinación de tecnologías que debe ser desarrollada por un grupo de colaboradores expertos que hagan de esta una herramienta innovadora para responder a las necesidades descritas en este documento. (Vargas y otros, 2020, p. 8).

Además, se planteó el *big data* como una posible solución «en los sistemas de gestión electrónica de archivo (SGDEA) de entidades públicas al aumentar la eficiencia en el manejo de la información y las comunicaciones» (Vargas y otros, 2020). Esta investigación realizada en Colombia plantea

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

que el *big data* es gestión de datos, a gran escala, es decir, la gestión y análisis de un volumen considerable de datos, y de gran variedad; por ello, el hecho de exponerlo en los SGDEA es relevante y acerca, de alguna forma, el concepto de gestión de datos al quehacer archivístico.

Una investigación más reciente, titulada *Saving Social Media Data: Understanding Data Management Practices* [Almacenamiento de datos de redes sociales: comprensión de las prácticas de gestión de datos], realizada por Libby Hemphill, directora asociada del Centro para la responsabilidad de las redes sociales; Margaret L. Hedstrom, profesora emérita de la Escuela de Información, y Susan Hautaniemi Leonard, investigador científico asociado (ICPSR) de la Universidad de Michigan, se enfocó en mostrar un primer acercamiento en torno a cómo las prácticas para el manejo y conservación de datos pueden adaptarse para los investigadores que los usan. En el desarrollo de la metodología se realizó un metaanálisis de artículos, cuyo propósito fue indagar cómo los investigadores usaban los datos de las redes sociales.

Para este fin se realizó la consulta en varias de estas plataformas sobre las interacciones, búsquedas y datos que usaban los investigadores. Se llegó a la conclusión de que los estándares para los datos de redes sociales no son suficientes y los archivos aún no apoyan este tema; además, «an archive that facilitates the preservation and reuse of social media data will need to contend with additional challenges in documenting data and its provenance, in describing what constitutes a “dataset” in this space, and in ensuring appropriate protections for personal and sensitive information» [un archivo que facilita la conservación y la reutilización de los datos de las redes sociales tendrá que enfrentarse a desafíos adicionales en documentar los datos y su procedencia, al describir lo que constituye un "conjunto de datos" en este espacio, y en garantizar las protecciones adecuadas para la información personal y confidencial] (Hemphill *et al.*, 2021, p. 24).

Cabe advertir que en el rastreo bibliográfico no se logró encontrar investigación alguna donde se relacione, específicamente, la gestión de datos con la archivística; sin embargo, se encontraron algunos avances y aproximaciones sobre la inmersión y el tratamiento de datos por parte de los archivistas, como la primera investigación sobre la formación y recuperación de información. Ahora bien, al hablar de documentos electrónicos se está haciendo referencia a datos no estructurados, de acuerdo con lo mencionado por Cunningham (2021), quien afirma: «en la medida en que se cree que los documentos encajan en el universo de los datos, generalmente se trata de datos no estructurados».

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Los correos electrónicos y los documentos de procesadores de texto son buenos ejemplos de datos no estructurados» (p. 33), los cuales son objeto de estudio de la archivística. Así mismo, en la función archivística se han desarrollado una serie de instrumentos; mediante estos son recolectados y usados los datos para un fin específico, entre otros, inventarios, tablas de valoración o retención documental, índices.

Con sano criterio, se han mostrado las investigaciones que, de algún modo, se relacionan con esta; no obstante, se vio la necesidad de allegar otros documentos de interés que tratan sobre la gestión de datos desde la perspectiva archivística. Este tema se ha tratado de manera parcial y ha estado presente en la gestión documental, tal como se evidencia en el trabajo «A la preservación de datos... ¡y más allá!», publicado en el año 2016 por Casellas & Serra (2016), responsable del Departamento de Gestión de Registros y delegado de Protección de Datos del Ayuntamiento de Girona (España), quien afirmó: «la preservación es una de las actividades propias de la gestión documental y está también se encuentra presente en la gestión de datos», lo que indica unos puntos de encuentro entre la gestión de datos y la gestión documental.

Otro de los artículos considerados, muestra el desarrollo de los procesos o actividades inmersas en la archivística, como la de extracción de datos, de acuerdo con lo expuesto por la profesora de Ciencias de la Información de la Universitat Oberta de Catalunya, García-Alsina, M., en su artículo «Transversalidad y gestión: documentos y datos al servicio de la toma de decisiones y la transparencia». En él menciona, como parte de los ámbitos de actuación de los archivos, la gobernanza de la información y la gestión de la calidad de los datos, y concluye:

los documentos además de tener un valor primario, la de actuar como evidencias de las actividades organizativas, tienen un valor secundario y que este valor solo se puede extraer mediante procesos de gestión de datos e información contenida en documentos ya sea en papel o en electrónico. (García, 2017, p. 14).

Finalmente, deja indicada casi que una dependencia entre los procesos, atribuyendo un tipo de relación; uno de los puntos clave de la presente investigación.

La segunda edición de la Guide to the Data Management Body of Knowledge (DAMA-DMBOK2), publicada en el año 2017, relevante para esta investigación en su actualización de la Global Data Management Community-DAMA International —una organización cuyo público

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

objetivo son los profesionales de datos para los cuales realiza el desarrollo de un cuerpo integral de estándares y prácticas de gestión de datos— se da teniendo en cuenta los avances que han ocurrido en torno a la gestión de datos, desde la publicación del primer ejemplar en el año 2009. La gobernanza de los datos, las herramientas tecnológicas y la posibilidad de realizar análisis *big data* son temáticas que llevaron a la publicación y actualización de esta guía.

En el año 2018, Bustelo Ruesta¹ plantea en su obra *Transformación Digital desde la perspectiva de la Gestión Documental* que se ha venido avanzando y evolucionando la gestión documental y lo plantea en tres momentos: a corto plazo, la gestión del documento electrónico, desde el punto de vista de su conservación y preservación; a mediano plazo, se menciona un avance sobre «la condición de fidedigno (y sus cuatro características) sólo puede alcanzarse cuando el documento tiene asociados sus correspondientes metadatos» (p. 7); y, por último, para el largo plazo, la autora plantea una posibilidad de comenzar a gestionar datos, documentos en XML, que permita la utilización de información de una manera más ágil; «sin embargo, por ahora parece haber una disociación entre la gestión de datos y gestión de documentos» (p. 9).

Otro de los escritos de Bustelo, «Gobernanza de datos y gestión de documentos», publicado en el 2021, hace referencia a un avance normativo, en países europeos, sobre la gestión de datos, donde dice:

se ha legislado profusamente sobre temas relacionados con los datos. Sin existir leyes específicas que abarquen la complejidad de la gestión de los datos se ha legislado sobre reutilización de los datos, protección de datos personales, datos abiertos del sector público y transparencia. (p. 71).

Un ejemplo relativo a esta normatividad es el Decreto de la Generalitat de Catalunya, que versa acerca de la gestión archivística de datos y activos digitales.

En ese mismo año, 2021, se hizo más concreta la preocupación por el tema de esta investigación, la cual ha sido abordada, entre muchos otros, como una transformación que va del documentalismo al dataísmo, tal como lo afirmó Manuela Moro-Cabero, profesora titular de la Universidad de Salamanca, adscrita al Departamento de Biblioteconomía y Documentación de la Facultad de Traducción y Documentación, en uno de los apartados de su artículo «Datos y archivos,

¹ Presidenta del AENOR/CNT50/SC1 (Comité Técnico Nacional de AENOR para Información y Documentación (CTN 50), Subcomité 1, Gestión de Documentos y aplicaciones).

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

retos para una profesión agitada por la transformación digital»; en este, además, hizo referencia a que deben gestionarse aquellos contenidos (bases de datos, blogs, webs, correos electrónicos, etc.), que no siendo documentos propiamente dichos, sí se derivan de procedimientos y contienen información con valores factibles para la organización.

Otro autor preocupado por el tema concreto es Jordi Serra, Director de Gobierno de la Información en el Consorcio de Educación de Barcelona. En 2021 realizó la publicación de varios textos, uno de ellos titulado «Los datos: ¿una amenaza? Gestión de documentos en la era de los datos», donde hace mención de la guía DAMA-DMBOK2 para la gestión de los datos y dentro del cual se lee que «la gestión documental ya existe referenciada como uno de los once ámbitos de la gestión de datos» (p. 57), cuando de la gestión de documentos y contenidos se trata. Igualmente, asegura que

aquello que realmente determina la calidad y el uso del dato, no es el dato en sí mismo, sino el ecosistema en el que se genera, tanto el organizacional como el tecnológico. Y esto es especialmente importante porque es sobre este ecosistema que puede actuar la gestión documental. (p. 51).

Para concluir este apartado, del mismo Serra (2021), *Gestión documental datificada*, obra en la que no solo menciona conceptos importantes, sino también la transformación, los cambios dados desde la archivística, y el llamado que se hace a los profesionales de archivo para que se involucren en la gestión de datos. En el texto se menciona que la calidad de los datos se garantiza si estos son fiables, consistentes y están disponibles, por último, se plantea la posibilidad de conseguir una gestión documental basada en datos, de seguir cinco pasos a saber: identificar, normalizar, integrar, servir y garantizar. Una de las conclusiones a las que llega Serra es en torno a la importancia de la gestión de datos para la toma de decisiones estratégicas, y que «en esta toma de decisiones realmente solo importan los datos y su fiabilidad; si los datos se autentican, el documento puede acabar convirtiéndose en una mera formalidad» (p. 525).

6. Metodología

Teniendo en cuenta el problema planteado sobre la producción masiva de datos y la necesidad de contar con estrategias para gestionarlos, así como la pregunta y objetivos propuestos para el desarrollo de la investigación, dentro de la metodología se definió el enfoque cualitativo, ya que

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

este es usado en la indagación que produce datos descriptivos, es decir, «orientada a describir con detalle y exhaustividad los fenómenos en uno o más momentos del tiempo» (Quecedo y Castaño, 2002). En este tipo de trabajo se asumió como método la investigación documental, misma con la cual se dio respuesta a la pregunta, partiendo de la búsqueda en material bibliográfico y referencial (normatividad, manuales, guías, estrategias, producción académica, entre otros), haciendo uso de la técnica revisión documental que, además, permitió observar el contexto donde se revela cómo se han desarrollado los procesos de gestión documental y datos.

En este sentido, el análisis del material consultado se realizó con el fin de describir cada concepto (gestión de datos y de documentos), así como cada uno de los elementos que lo compone, es decir, no solo sus definiciones, sino también los tipos de datos y documentos, los principios y los procesos tanto estratégicos y transversales como los técnicos.

Posteriormente, se dieron a conocer las diferencias y similitudes entre la gestión de datos y la gestión documental, dando cumplimiento al segundo objetivo definido. Si bien se compararon varios elementos, cabe resaltar que luego de encontrar en la bibliografía categorías similares entre ambos, fue viable realizar este comparativo; no obstante, el desarrollo del segundo objetivo no responde a ninguna metodología previamente establecida; por tanto, se tomó cada uno de los elementos y se compararon con el fin de hallar similitudes, diferencias y puntos de complementariedad.

Finalmente, el conocimiento de la gestión documental y la gestión de datos, obtenido luego de describir y comparar cada uno de los elementos (definiciones, tipos de datos y documentos, principios y procesos), posibilitó definir las estrategias para la gestión archivística de datos en Colombia.

6.1 Recolección de información

Durante el proceso de recolección, se procedió mediante el análisis de documentos —normatividad, manuales, guías, estrategias, producción académica, entre otros— de acuerdo con las categorías definidas para la investigación sobre los conceptos gestión de datos y gestión de documentos, más específicamente, sobre el ciclo de vida de los datos y documentos; así mismo, la necesidad de gestionar la información y la archivística, y los datos; para hacerlo, «no hay categorías prefijadas,

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

se extraen del texto (en registro escrito u oral) patrones de conducta o situaciones creando unidades inferidas a partir del registro narrativo» (Quecedo & Castaño, 2002, p. 23).

En el proceso de recolección se usaron dos instrumentos: una matriz bibliográfica, diseñada para llevar el control de toda la bibliografía consultada, en la cual, a su vez, para cada documento encontrado se registraron los datos de autor, año de publicación, tipo de documento, idioma, editorial, nombre o título, país, y URL. Esta matriz facilitó, de manera general, la identificación de los textos, objeto de análisis, con el propósito de revisar si fuera necesario o no ser usados durante la investigación; adicionalmente, se usaron fichas de contenido en donde se detallaron los textos usados en el desarrollo de la investigación y, lógicamente, se registraron campos como: título, autor, palabras claves, resumen, cita y ubicación/URL. Finalmente, se presenta la cantidad de bibliografía consultada, las ecuaciones usadas y los criterios de selección y/o exclusión que se tuvieron al momento de revisar la documentación:

Tabla 1: Cantidad de bibliografía

BIBLIOGRAFÍA REVISADA	CANTIDAD
Artículo	69
Guía	9
Informe	1
Libro	1
Manual	3
Normatividad	11
Política	1
Conferencia	2
TOTAL	97

Fuente: Elaboración propia.

*De estos, 91 se encuentran en español, 4 en inglés y 2 en portugués.

Ahora bien, para la obtención de esta bibliografía se usaron las siguientes ecuaciones:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Tabla 2: Ecuaciones para la búsqueda de bibliografía

ECUACIONES	
Gestión de datos AND archivística AND Colombia	Data management AND archival science AND Colombia
Gestión de datos AND gestión documental	Data management AND records management
Archivística AND datos	Archival science AND data
“gestión documental datificada”	“records datafied management”
Big data AND archivística	Big data AND archival science

Fuente: Elaboración propia.

Finalmente, se aplicaron los siguientes criterios de inclusión y/o exclusión:

Tabla 3: Criterios de inclusión y/o exclusión

CRITERIOS DE INCLUSIÓN	CRITERIOS DE EXCLUSIÓN
Metodología: Afinidad metodológica con la investigación	Enfoque: Bibliografía cuyo alcance no fuese de interés, es decir, los datos vistos o tratados desde otras áreas como la estadística, matemática, etc.
Idioma: Publicaciones en español e inglés.	Vigencia y temporalidad: Bibliografía publicada cuyos conceptos no fuesen de interés en la actualidad.
Temática: Afinidad temática con la investigación	Fuentes: Bibliografía encontrada en bases de datos diferentes a Scielo, ScienceDirect, Scopus, Springerlink, Google Scholar y Redalyc.

Fuente: Elaboración propia.

6.2 Análisis y presentación de información

Teniendo en cuenta que «la finalidad del análisis es una mayor comprensión de la realidad» (Quecedo & Castaño, 2002, p. 25), este se realizó de modo alterno a la recolección de datos, y más fuertemente finalizado el proceso de recolección, lo cual favoreció la adquisición de conocimiento suficiente para manejar el tema de investigación y lograr cada objetivo trazado. Para el análisis de la información se usaron ilustraciones que permitieron graficar los procesos de gestión de datos y

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

gestión documental; a través de ello, se lograron identificar todos los procesos que los conforman, manera como se brindó una síntesis de las características de cada uno y una mejor comprensión de los procesos y su articulación. Además, se ilustraron los principios, consolidándolos en una sola imagen para cada caso; así mismo, a modo de mapa mental, se ilustró la comparación realizada entre los conceptos de gestión de datos y gestión documental, brindando una ayuda visual para una mejor comprensión de las relaciones que existe, entre estos.

Por otra parte, se usaron tablas, una serie de cuadros donde se muestran los comparativos realizados, ya que para un mejor entendimiento se presentan comparaciones entre cada elemento de la gestión documental y la gestión de datos, es decir, los principios, los tipos y los procesos —tanto estratégicos como técnicos—. Lo dicho, consintió reflejar la presencia de cada ítem en la gestión de datos y la gestión documental, según correspondiera. Ya en el caso de los principios se precisó la necesidad de mostrar dos tablas, una donde se observaron las similitudes entre los principios; otra, donde se verificó si el principio estaba relacionado con algún elemento de la gestión documental o la gestión de datos, según aplicara.

Adviértase, pues, que en esta investigación se desarrollaron y compararon conceptos, se definieron posibilidades de complementariedad, se identificaron vacíos en la gestión de datos, en relación con la gestión documental, y se presentaron las estrategias para la gestión archivística de los datos en Colombia, formadas por componentes y estos, a su vez, por actividades.

Finalmente, una vez analizada y consolidada toda la información, se hizo uso de una estructura jerárquica donde se establecieron las estrategias, discriminando sus componentes y actividades, tomando como referencia el Modelo de Gestión Documental y Administración de Archivos (MGDA), publicado por el Archivo General de la Nación en el año 2020.

7. Resultados

Como producto de la revisión bibliográfica realizada en textos académicos, artículos de investigación, guías y manuales, se realizó la caracterización de la gestión de datos y la gestión documental. Se optó por desarrollar su definición, tipos de datos y documentos, sus principios y los procesos que los componen, como características relevantes para esta investigación, que sirvieron de base e información fundamental para la ejecución de los objetivos planteados, pues, de acuerdo

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

con lo expuesto en la problemática que se pretende tratar, se hace necesario comprender la gestión de datos y la gestión documental de forma conceptual, teniendo en cuenta la relevancia de la gestión de datos como una oportunidad dentro de la disciplina archivística y la aplicabilidad actual que esto tiene, pues cada vez se producen y consumen un gran volumen de datos en las organización que necesitan no solo ser explotados, sino también gestionarlos, administrarlos de forma integral; de esta manera fue de vital importancia comprender conceptualmente la gestión de datos y la gestión documental, a pesar de las dificultades presentadas por la falta de bibliográfica específica sobre el tema de investigación, la gestión de datos desde la perspectiva archivística.

En vista de esta situación se consideró pertinente obtener el conocimiento necesario sobre el tema a trabajar y así poder hacer la comparación de conceptos, mediante la elaboración de un cuadro comparativo, a partir del cual se analizaron en profundidad las características de la gestión de datos y la gestión documental, desde una óptica de similitudes, diferencias y posibilidades de complementación. Finalmente, se plantearon algunas estrategias de gestión archivística de datos en Colombia.

7.1 Características de la gestión de datos y la gestión documental.

Desarrollar las características de la gestión de datos y de la gestión documental, de manera independiente, fue fundamental para comprender aquellos conceptos relacionados y las particularidades de cada uno, a la par que formular las estrategias de gestión de datos desde una perspectiva archivística. Para la gestión de datos se tomaron como referencia los planteamientos del DAMA BOOK de la Global Data Management Community y del Departamento Nacional de Planeación y el Ministerio de TIC de Colombia; para el caso de la gestión de documentos, se tomaron como referencias las NTC-ISO 30300 y 15489, las publicaciones de los investigadores Serra Serra (2021), Bustelo Ruesta (2018 y 2021), Cruz Mundet (2011) y Heredia Herrera (2018 y 2019), el Manual para la Implementación de un Programa de Gestión Documental, el Decreto 1080 de 2015 del Ministerio de Cultura de Colombia, además de publicaciones realizadas por el Archivo General de la Nación, como el manual para la implementación de un Programa de Gestión documental.

Las características tenidas en cuenta para el desarrollo de este apartado fueron: la definición del concepto a tratar, los principios, los tipos y los procesos que conforman la gestión de datos y de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

documentos. Las características permitieron comprender y tener mayor claridad sobre ambos conceptos y los aspectos relevantes de cada uno de ellos.

7.1.1 La gestión de datos

La *gestión de datos* se entiende como la ejecución de una serie de acciones y procesos, consolidados en planes, políticas, programas y estrategias que, correctamente aplicadas y evaluadas, permiten mantener un control de los datos durante su ciclo de vida, un procesamiento técnico y operativo que comprende la extracción, la clasificación, la limpieza y posterior visualización y uso de los datos. Estos procesos también buscan mantener y garantizar la calidad en el activo de información, de manera que se acceda a datos completos y fiables, que cumplan con los requerimientos y las necesidades de cada solicitante o usuario.

En este orden de ideas, y para el caso colombiano, se cuenta con el documento Conpes del Departamento Nacional de Planeación sobre la Política Nacional de Explotación de Datos, con el cual se busca promover el aprovechamiento de los datos, el cual se ha estado enfocando, principalmente, en la analítica, entendiendo este proceso como el uso y análisis de los datos, proceso en el que se extrae información relevante para alguna necesidad expresa en la entidad; de esta manera es como se evidencia que se está dejando a un lado la gestión y el tratamiento de los mismos, es decir, no se tienen en cuenta los procesos por los cuales atraviesan los datos desde su recolección o generación hasta su posterior conservación. Se está en un punto medio donde su explotación, la importancia que revisten para la toma de decisiones y su disposición, en cumplimiento de las directrices de acceso abierto o la publicación de activos de información, cobran mayor trascendencia que la misma arquitectura de los datos, la calidad, el soporte de las operaciones, la gestión de datos y contenido (datos no estructurados) y la interoperabilidad.

Una última consideración en este apartado. Otro de los enfoques que se ha ido desarrollando es la gobernanza de datos, tal como se evidencia con la guía publicada por el Ministerio de TIC, donde se brindan lineamientos para que en las entidades públicas se logre tener el control y administración integral de los datos.

7.1.1.1 Tipos de datos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Existen diversos tipos de datos, por ejemplo, aquellos simples (números enteros, decimales, texto), compuestos (tablas, listados, diccionario), transaccionales (datos de programas orientados a eventos comerciales y empresariales), maestros (datos definidos en una matriz que buscan dar orientación y contexto, necesarios para comprender las transacciones), entre otros. Hacer esta categorización es relevante para gestionarlos, ya que de acuerdo con la condición de cada tipo de dato, se tienen ciertas especificaciones propias para su uso y publicación en Colombia, como el caso de los datos públicos que son de libre acceso, a diferencia de los datos personales que requieren «el consentimiento previo del titular para su almacenamiento, compartición y uso» (Departamento Nacional de Planeación, 2018), según los «derechos a la intimidad, privacidad, la garantía del secreto profesional, comercial e industrial» (Departamento Nacional de Planeación, 2018, p. 26).

Teniendo en cuenta la variabilidad que presentan los datos y la necesidad de tener conocimientos del tipo que se pretende gestionar, para que este proceso sea más eficaz y se puedan brindar las estrategias acordes para cada uno, se toman los siguientes tipos de datos que engloban los expuestos anteriormente:

- *Datos estructurados*: agrupados en bases de datos, de manera que se identifica plenamente la relación entre estos y que se cumpla con el modelado de los mismos.
- *Datos no estructurados*: contrario de los datos estructurados, estos no cumplen con el modelado; «en esta categoría se incluyen, por ejemplo, las imágenes, texto, audios, contenidos de redes sociales, vídeos» (Departamento Nacional de Planeación, 2018, p. 25).
- *Datos semiestructurados*: la organización, agrupación y representación de estos datos, no cumple con el modelado de datos; además, no se tienen definidas sus relaciones; «en esta categoría se incluyen contenidos de e-mails, tweets, archivos XML» (Departamento Nacional de Planeación, 2018, p. 25).

Otras de las divisiones que puede hacerse respecto a los tipos de datos, se explica de acuerdo con la fuente de la cual fueron extraídos; de ahí derivan los *datos públicos*, generados o recolectados por las entidades públicas; los *datos privados*, pertenecientes a una empresa particular; así mismo, los *datos impersonales* y *personales*, referentes o no a una persona y su derecho a la intimidad.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

7.1.1.2 Principios de la gestión de datos

Con estos se busca brindar orientaciones básicas respecto a la gestión de datos. Para este fin se tomaron aspectos establecidos por la Comunidad Global de Gestión de Datos (The Global Data Management Community) en su guía DAMA-DMBOK2 y en el documento Conpes para la explotación de datos del Departamento Nacional de Planeación, donde se hallaron conceptos que sirvieron para complementar los principios enunciados en la guía. Cabe resaltar que aquellos que se enuncian a continuación, son producto de un análisis realizado a los referentes mencionadas anteriormente; en este sentido, fueron combinados, suprimidos e incluidos en otros, con el fin de brindar mayor claridad frente a este tema y evitar la redundancia en la información presentada. Los principios de la gestión de datos son (figura 1):

Figura 1 Principios de la gestión de datos



Fuente: Elaboración propia

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

A continuación, se describen los principios de la gestión de datos, enunciados en la figura 1:

- *Los datos son un activo con propiedades únicas*, ya que estos no se consumen cuando son usados; además, cobran valor en la medida en que se usan, es decir, cuando se agrupan y se extrae información de ellos. Estos activos no tienen límite de uso, en razón a que pueden ser empleados para diferentes finalidades, de acuerdo con lo requerido por la entidad; sin embargo, para determinar su valor, no se cuenta con una fórmula universal y es por ello que cada entidad, de acuerdo con sus características, puede determinar el valor cualitativo y cuantitativo de sus datos, dando paso a otro de los principios enunciados por la Comunidad Global de Gestión de Datos sobre la posibilidad de que el valor de los datos puede y debe expresarse en términos económicos.
- *La gestión de datos es la gestión del ciclo de vida*: los datos atraviesan por unas etapas consideradas ciclo de vida, que va desde su producción o extracción hasta su análisis y uso; por esta razón, se dice que la gestión de datos es la gestión del ciclo de vida de estos, ya que son actividades que van a garantizar y a definir su trayectoria. Cabe resaltar que este ciclo es más complejo que el de un documento, ya que los datos poseen unas características diferentes; además, se verá guiado por el tipo de dato y sus necesidades específicas durante toda su gestión.
- *Gestionar datos significa gestionar la calidad de los datos*, si se tiene en cuenta que durante este proceso se contemplan aspectos que permiten que los datos sean recolectados, generados, almacenados y dispuestos de una forma regulada, lo que facilita que estos cumplan ciertos requerimientos como ser usables, fiables y estar completos. Se consideran datos de calidad, y lo que se está realizando a lo largo de su gestión, durante todo su ciclo de vida, es garantizar su calidad. Por esta razón, se tiene este principio, pues gestionar correctamente los datos, es garantizar su calidad.
- *Se necesitan metadatos para administrar los datos*, máxime si no se pierde de vista que los metadatos permiten la identificación de una serie de características propias de los datos, mientras brindan información sobre su contenido y estructura, así como la identificación de un contexto. Una matriz de esta naturaleza es necesaria para determinar una normalización

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

de los datos, por cuanto posibilita, de una parte, tener un control sobre los datos y, de otra, planificar la gestión de datos desde su producción o recolección. Lo dicho conduce a concluir que el normalizar la producción de los datos permite la usabilidad de estos.

- *La gestión de datos es transversal e interdisciplinaria:* involucra todas las áreas de la entidad, desde donde los datos son recolectados o generados; además, se requiere de una amplia gama de profesionales, con diferentes habilidades, perspectivas y experiencias, debido a que los diferentes tipos de datos representan necesidades específicas en su forma de gestionarlos. Teniendo en cuenta que vienen de diferentes áreas, temáticas, formas, estructuras y serán usados en un entorno tecnológico, con diferentes finalidades, es necesaria la participación de diferentes áreas y profesionales, para tener la posibilidad en la organización de gestionar, usar y tomar decisiones con los diferentes datos que se tienen.
- *La gestión de datos requiere una perspectiva integral:* de manera que se implementen en toda la entidad las actividades que la conforman, mediante una cadena de custodia definida para el manejo de los datos, de acuerdo con el tipo; en este sentido, es posible garantizar la calidad de los datos y dar cumplimiento a uno de los procesos estratégicos: la gobernanza de los datos, definido en la guía para el gobierno de datos publicada por el Ministerio de TIC de Colombia como «el ámbito enfocado en identificar los custodios y consumidores del dato, y en ejecutar los requisitos relacionados de cumplimiento y responsabilidad» (Ministerio de Tecnologías de la Información y las Comunicaciones, 2019).
- *La gestión de datos incluye su gestión de riesgos:* la pérdida, robo o mal uso de los datos, conlleva a una serie de implicaciones éticas y legales a la hora de usarlos; por tanto, para la gestión de riesgos se hace necesario elaborar una matriz en la que se deje registro de los peligros a los que están expuestos los datos y, por ende, se planteen alternativas o actividades que permitan mitigarlos o atenderlos cuando se materialicen.
- *Los requisitos de gestión de datos deben impulsar las decisiones de tecnología de la información:* la gestión de los datos y las tecnologías de la información están estrechamente relacionadas, ya que son instrumentos que favorecen la extracción, generación, gestión y almacenamiento de los datos. Se busca que al usar las tecnologías de la información se

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

fortalezca la gestión de los datos, se conciba la tecnología como una oportunidad y no como una necesidad en la entidad.

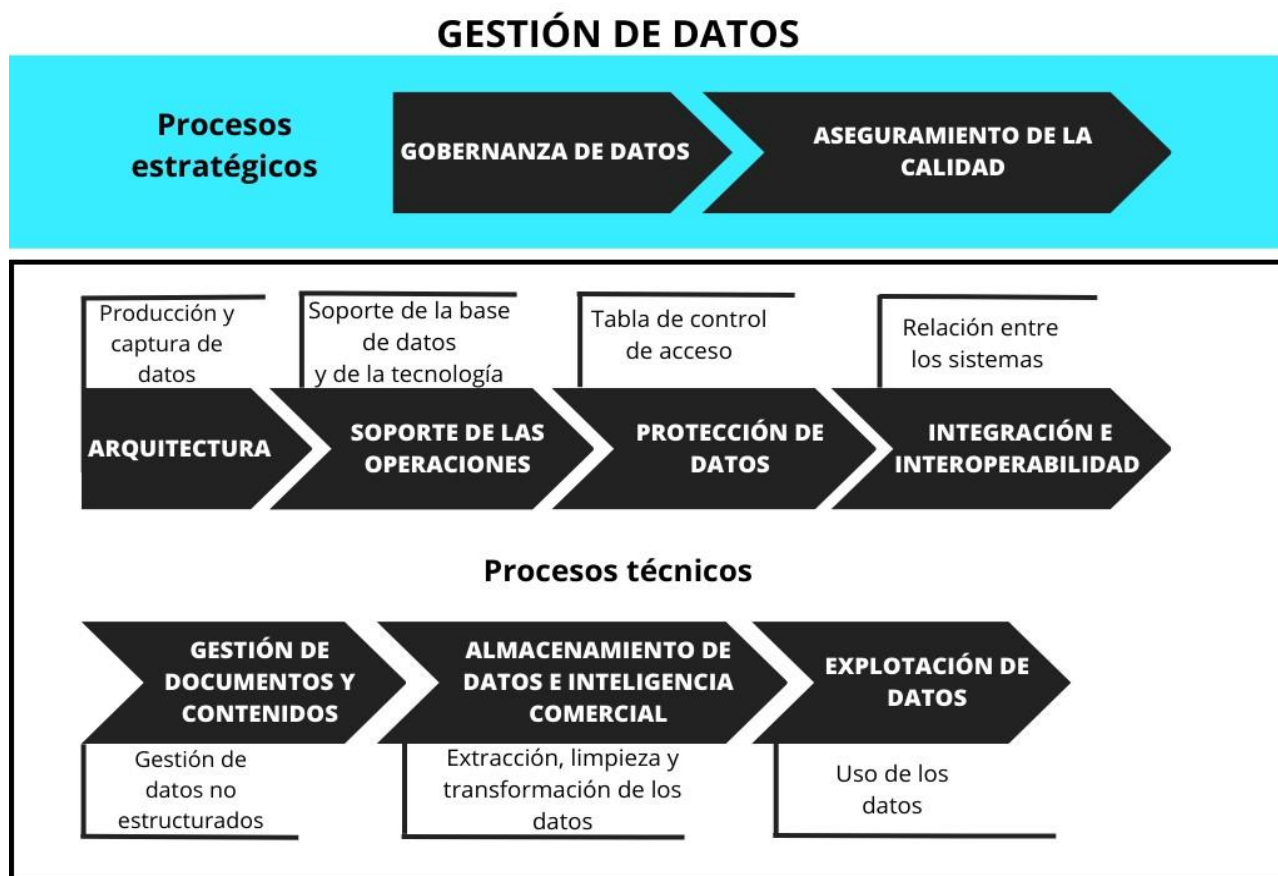
7.1.1.3 Procesos de la gestión de datos

Para plantear los procesos de la gestión de datos, se tuvo en cuenta lo expuesto por el Global Data Management Community, en la guía DAMA-DMBOK2, porque en este documento se desarrollan, de manera integral, cada uno de los procesos. Ahora bien, como referente colombiano, fueron tenidos en cuenta el documento Conpes para la explotación de datos del Departamento Nacional de Planeación y la guía para el gobierno del dato del Ministerio de TIC, relacionados con el tratamiento de los datos, donde se plantean actividades para garantizar su aprovechamiento y difusión.

Para iniciar, en primer lugar, se presentan los procesos, tanto estratégicos como secuenciales y, posteriormente, se desarrolla cada uno, teniendo en cuenta que estos son una guía para la gestión de datos:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Figura 2 Procesos de la gestión de datos



Fuente: Elaboración propia

- *Procesos estratégicos de la gestión de datos:* como lo indica Rincón (1998), estos procesos «facilitan las guías de actuación para los procesos claves y están basados en la escucha de la voz del cliente o del mercado». Implica la definición de una serie de lineamientos, orientando cada proceso de la gestión de datos. A continuación, se enuncian y detallan los procesos de gobernanza y calidad de los datos, como procesos estratégicos de la gestión de datos:

A. Gobernanza de datos: Este proceso tiene como objetivo mantener control sobre los datos durante todo su ciclo de vida. Para el caso de Colombia, el Ministerio de las TIC (2019) ha publicado una política al respecto, donde se plantea que la gobernanza «es el ámbito enfocado en identificar los custodios y consumidores del dato, y en ejecutar los requisitos relacionados de cumplimiento y responsabilidad» (p. 14); esta gobernanza «se expresa a través de una política que articula los controles, las metas de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

negocio, objetivos estratégicos, los procesos involucrados e indicadores de gestión» (Departamento Nacional de Planeación, 2018, p. 29), proceso que se ocupa de identificar el origen de la producción de los datos, así como los responsables de estos a lo largo de su ciclo de vida, siendo este, a su vez, el fin último de la gestión de datos, el cual es tener control para la toma de decisiones estratégicas en la entidad.

Para lograr la gobernanza de datos, la política para el gobierno de los datos del Ministerio TIC plantea la ejecución de cuatro fases, enunciadas a continuación:

Primera fase: Enfocada en identificar los datos que se consideran críticos, entendiendo estos como aquellos producidos y recopilados por la entidad que dan cuenta de su gestión diaria y que por ende son de gran utilidad para la entidad como activos de información, y que por esta razón deben ser tratados o gestionados.

Segunda fase: En esta cobra gran relevancia la gestión de metadatos, ya que se hace necesario la elaboración de una matriz que permita caracterizar los datos que serán extraídos o generados y su estructura; de esta manera se busca una mejor administración durante su ciclo de vida, abarcando la arquitectura de los datos y los responsables de estos.

Tercera fase: En esta se aprovechan los insumos generados en las primeras fases, mediante la elaboración de un diccionario donde se definan los datos críticos, una matriz de metadatos y la definición de datos maestros, con el fin de manejar un lenguaje normalizado en la entidad y así lograr un entendimiento en común.

Cuarta fase: Plantea la elaboración de la política de gobernanza de datos, partiendo del diagnóstico que representan las tres primeras fases; esta política deberá definir responsables, actividades a ejecutar, metas para garantizar la calidad de los datos e indicadores de seguimiento.

La gobernanza de datos tendrá éxito si se garantizan datos disponibles, de calidad, reutilizables e interoperables, de manera que el acceso a los datos sea un servicio para la entidad, sus proveedores, usuarios o las personas que lo requieran y puedan o tengan los permisos para acceder a estos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

B. Aseguramiento de la calidad de los datos: está presente durante todo el ciclo de vida de los datos. Los requisitos de calidad de los procesos dependen de las necesidades del usuario o la entidad, para que los datos sean fiables y confiables; sin embargo, estos pueden ir cambiando con el tiempo, porque es una actividad constante, pues durante todo su ciclo de vida se está en búsqueda o se pretende garantizar datos exactos, que estén completos, que no se encuentren duplicados, que sean íntegros y consistentes.

Es recomendable implementar herramientas que permitan medir la calidad de los datos, con indicadores y acciones de mejora definidas en una política, acorde con el modelo de datos y la formulación de datos maestros y de referencia, que buscan no solo una normalización, sino también garantizar la calidad y evitar repeticiones. Este proceso busca garantizar que se cuente con datos aptos para el consumo, es decir, que estos sean íntegros, correctos, consistentes, actualizados y que cumplan con los demás requerimientos que permitan la toma de decisiones; además, para garantizar su calidad es necesario contar con instrumentos, por ejemplo, indicadores, que permitan mantener un control sobre las características óptimas de los datos, la implementación del ciclo PHVA, la elaboración de diagramas donde se puedan identificar problemas o falencias relacionados con los datos, entre otros.

- *Procesos técnicos de la gestión de datos:* Además de los procesos estratégicos mencionados para la gestión de datos, a continuación, se presentan siete procesos secuenciales mediante los cuales se puede lograr una articulación y gestión integral de los mismos:

A. Arquitectura de datos: en este se contempla la construcción, generación, captura y extracción de los datos, de manera que estos se produzcan de acuerdo con las funciones de la entidad, es decir, con criterios establecidos. En este sentido, la arquitectura permite definir qué datos serán generados o recopilados, cómo será el proceso de recolección o captura, en qué sistemas se almacenarán y cómo serán agrupados, en los casos que así corresponda. Cabe resaltar que dentro de este proceso se encuentra el modelado de datos, que consiste en una serie de requisitos que deben ser tenidos en cuenta a la hora de producirlos y recopilarlos en bases de datos, de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

forma tal que se establezcan aquellos que serán manejados, las relaciones entre estos y sus características. Por ejemplo, en una base de datos donde se relacionan los pedidos a proveedores, se debe definir cuáles conformarán la base de datos (nombre del proveedor, producto solicitado, costo, entre otros) y las relaciones entre estos.

La captura es otro elemento importante a considerar dentro de la arquitectura de datos, la cual hace referencia al momento en que «son extraídos los datos, ya sea de forma tradicional, es decir, mediante formularios, encuestas, trámites o de forma digital, usando los dispositivos de captura como lo son sensores» (Departamento Nacional de Planeación, 2018). En este mismo sentido, se habla de la generación y recolección, como la «creación de los datos digitales por personas como resultado de la interacción con sistemas, herramientas y servicios digitales, o por máquinas con programas de software y dispositivos de hardware que capturan fenómenos» (Departamento Nacional de Planeación, 2018).

Esta captura o generación de datos se debe realizar, de una parte, de forma manual, con el uso de herramientas tecnológicas que permitan diligenciar encuestas en aplicaciones web, por ejemplo, que posibilite la búsqueda y uso de los datos de una manera eficaz; de otro, de forma automática, en el evento de aquellos registros de datos generados en un sistema por medio de sensores —como las cámaras que registran el ingreso de una persona a una entidad—. Cabe resaltar que en ambos casos se debe contar con una definición previa de los datos que serán capturados o recolectados.

- B. Soporte a las operaciones de datos:** corresponde al soporte y trámite de los datos. Dentro de este proceso se encuentra el mantenimiento de las bases de datos, enfocado en los procesos de *software*. Así mismo, se tiene el soporte de las bases de datos, donde se contempla el mantenimiento de los sistemas donde estas son gestionadas. Los sistemas de información para ejecutar operaciones o transacciones son muy comunes; actualmente las entidades confían en estos para la ejecución de funciones. Las operaciones de datos se realizan en los sistemas organizacionales, razón por la cual se garantiza la continuidad del negocio. Sin embargo, cuando un sistema falla, la operación de la entidad se ve afectada, generando retrasos e, incluso, interrupciones

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

completas en su funcionamiento. Por ello, es deseable y crucial contar con un buen soporte desde el área de las TIC.

Teniendo en cuenta las necesidades propias de cada entidad, los datos que maneja, las transacciones, los clientes o usuarios, y su finalidad de negocio, se podrá determinar la tecnología necesaria para la gestión de sus datos. Este proceso requiere de un monitoreo constante, que facilite el mantenimiento de la tecnología, así como la generación de copias de seguridad, para de esta manera prevenir fallas en el sistema y pérdida de información.

- C. *Protección de datos*: Se enfoca en garantizar la seguridad de los datos, teniendo en cuenta los tipos que maneja cada organización; se tiene entonces que los datos personales tienen algunos niveles de reserva y, dependiendo del objeto o las funciones de la entidad, los datos que se capturan o generan en ella serán de carácter público o privado.

En el caso particular de Colombia, está la Ley 1581 de 2012, que tiene por objeto

desarrollar el derecho constitucional que tienen todas las personas a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bases de datos o archivos, y los demás derechos, libertades y garantías constitucionales a que se refiere el artículo 15 de la Constitución Política; así como el derecho a la información consagrado en el artículo 20 de la misma.

Valga anotar que regula la protección de los datos personales, por tanto, hoy en día es muy común encontrar una serie de condiciones para el tratamiento de los datos, autorizaciones, políticas y demás herramientas implementadas por cada entidad, para dar a conocer a sus usuarios, trabajadores y proveedores el uso que se dará a ciertos datos. Más recientemente con la Ley 1712 de 2014 que tiene por objeto «regular el derecho de acceso a la información pública, los procedimientos para el ejercicio y garantía del derecho y las excepciones a la publicidad de información». Tiene como fin la protección de datos personales y el índice de información clasificada y reservada.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Para determinar la ruta que permita garantizar la protección de los datos, se pueden presentar una serie de elementos que, según lo establecido por la Global Data Management Community en la DAMA-DMBOK2, se encuentran orientados a la definición de las partes interesadas, de manera que se conozcan las necesidades de privacidad y confidencialidad de las partes implicadas (clientes, usuarios, estudiantes, ciudadanos, entre otros), las regulaciones gubernamentales, para tener en cuenta las habilitaciones o restricciones legales que se tienen para el acceso a los datos, de acuerdo con su tipo, su productor y personas involucradas. Otros de los elementos que deben tenerse en cuenta son las preocupaciones comerciales de propiedad; es decir, las bases de datos con que cuenta una entidad y que son de su propiedad, como son aquellas donde se registran sus proveedores, sus materias primas y demás aspectos fundamentales para su funcionamiento. También se tienen las necesidades de acceso legítimo, donde se plantea un control, se definen roles y su nivel para consultar ciertos datos y, finalmente, unas obligaciones contractuales, los acuerdos entre los involucrados contractualmente (personas, compañías, entre otros) para el manejo de los datos entre las partes y de cara a los interesados en estos.

- D.** *Integración de datos e interoperabilidad:* teniendo en cuenta el uso de herramientas tecnológicas y sistemas informáticos para la gestión y almacenamiento de los datos, se hace referencia a la integración e interoperabilidad de estos en los almacenes o repositorios y las aplicaciones, de forma que los sistemas puedan relacionarse, comunicarse entre sí o ejecutar tareas en conjunto.

Dentro de las actividades a ejecutar como parte de este proceso, la Global Data Management Community en la DAMA-DMBOK2 plantea:

- a) La migración y conversión de datos, entendida como la actividad de moverlos de un sistema a otro, de un formato a otro y transformarlos para que estos sean compatibles con la estructura del almacén de destino.
- b) La consolidación de datos, de manera que se agrupen para obtener bases de datos más completas.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- c) El intercambio de datos entre aplicaciones y diferentes sistemas para compartir y realizar transacciones; es decir, la distribución de datos entre almacenes y centros de datos, garantizando el intercambio en diferentes bases de datos y unidades de almacenamiento.
- d) Archivar datos, almacenarlos en un formato legible y accesible.
- e) La gestión de interfaces de datos, para definir el origen de estos, examinarlos, tener presente el esquema de datos y crear correlaciones para transformar e importarlos para su uso.
- f) Obtención e ingesta de datos externos, extracción y toma de los mismos de otro sistema, incluso, ajeno a la entidad.
- g) La integración de datos estructurados y no estructurados, con la posibilidad de gestionarlos de una manera conjunta.

Se aspira a que los datos puedan ser compartidos, acceder a ellos sin mayor dificultad y garantizar su uso a lo largo del tiempo; es por ello que se debe tener en cuenta el formato donde se encuentren, para que estos sean procesados, almacenados y visualizados en cualquier sistema. En este proceso cobran gran relevancia los metadatos propios de los datos y del sistema, como parte esencial de una correcta interoperabilidad; es importante planificar este proceso, y validar si la relación entre sistemas o bases de datos puede llevarse a cabo sin ningún inconveniente. Así mismo, determinar si este proceso se realizará con datos o con bases de datos adquiridas, para lograr tener control y definir las necesidades para cada grupo de datos, de modo que se prevenga la pérdida de los mismos.

- E.** *Gestión de documentos y contenidos:* contempla la gestión de los datos no estructurados como son textos, imágenes, correos electrónicos, entre otros, con el fin de garantizar su calidad y seguridad. Este se realiza de acuerdo con el ciclo de vida de los documentos, desde la producción o recepción, hasta su eliminación o conservación; así es como se plantea un enfoque, dentro de la gestión de datos, para la gestión documental; haciendo referencia a las políticas que tienen los sistemas, así como la identificación de escritos vitales y el cronograma de conservación.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Dentro de este proceso se plantean una serie de actividades que van desde la planeación del ciclo de vida de los documentos, donde se hace mención que estos son creados o recibidos, distribuidos, almacenados y, posiblemente, destruidos, al igual que se hace en la gestión documental desde la teoría archivística. En esta primera actividad de planeación, se hace uso de sistemas de clasificación, indización y taxonomías para el almacenamiento y recuperación de los documentos, manteniendo una descripción con el uso de los metadatos necesarios para garantizar el ágil acceso o consulta de los datos. Se plantea, entonces, la formulación de una política, donde se definan los responsables de la gestión de documentos, así como los procesos de organización de estos, protección de los registros vitales y métodos adecuados para su destrucción.

- F.** *Almacenamiento de datos e inteligencia comercial:* este proceso busca reunir todos los datos en un mismo repositorio, para lo cual es necesario tener en cuenta varios aspectos: las condiciones en que se encuentran los datos a almacenar, además de atender a las personas involucradas en el proceso, de manera que se identifique de dónde vienen los datos, para qué y cuándo serán usados, y garantizar que el almacén de datos permita la ejecución de las actividades y las transacciones, así como la visualización de los datos para su posterior uso.

En este proceso se contemplan la extracción, limpieza y transformación. La extracción es el proceso mediante el cual son tomados los datos y arrojados en un sistema; la limpieza, por su parte, es una actividad que consiste en darles un formato normalizado, de acuerdo con lo establecido para ello, así como detectar campos vacíos en una base de datos, determinar cómo proceder e identificar los que están duplicados; dicho de otro modo, se trata de preparar la base para el procesamiento de información. Finalmente, la transformación entraña que los datos puedan ser usados para la toma de decisiones y que estos se puedan presentar a las personas interesadas. Estos procesos de limpieza y transformación se hacen para los datos estructurados, aquellos definidos en un modelo.

Una vez se definen los datos que irán al almacén y los perfiles de consulta donde se determina quiénes son los responsables de estos, se pasan los datos al sistema. Las

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

indicaciones para realizar este proceso deberán estar al alcance de las personas comprometidas, personal de la entidad y clientes o usuarios que puedan tener acceso, de acuerdo con las necesidades identificadas en cada uno.

- G. *Explotación de datos:*** en este proceso se busca generar valor social y económico con el uso de los datos, garantizando que se obtengan nuevos bienes, servicios o procesos de innovación basados en datos. Esto significa tener pleno conocimiento del uso que se dará a los datos capturados o recolectados, con el fin de generar servicios de información y facilitar la toma de decisiones en las organizaciones. En Colombia se cuenta con el documento Conpes sobre explotación de datos, en el cual se afirma que este proceso busca analizar los datos e identificar en estos, «hechos, relaciones, patrones ocultos de comportamiento de variables, correlaciones y tendencias» (Departamento Nacional de Planeación, 2018, p. 29). Se busca describir, diagnosticar e, igualmente, predecir, situaciones mediante el análisis de los datos disponibles.

Este análisis se realiza con herramientas tecnológicas, programas que faciliten este proceso mediante la ejecución de parámetros previamente establecidos, de acuerdo con las necesidades de cada entidad y la finalidad de la explotación de estos datos, donde la visualización cobra relevancia, pues en esta se brinda una mejor comprensión del análisis realizado, tanto para el responsable como para su usuario o público.

7.1.2 La gestión documental

La gestión documental es el «conjunto de normas, técnicas y conocimientos aplicados al tratamiento de los documentos desde su diseño hasta su conservación permanente» (Mundet, 2011, p. 18); comprende una serie de procesos que abarcan el ciclo vital del documento, que va desde la planeación hasta la disposición final de los documentos. Durante su ejecución, se van teniendo en cuenta una serie de requisitos técnicos, normativos y tecnológicos que facilitan que los documentos sean gestionados, almacenados y conservados, a partir de las necesidades o realidades particulares de cada entidad.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

En esta línea de sentido, se inicia con la generación y recepción de un documento, su gestión, la posterior agrupación de estos en unidades documentales simples (con características homogéneas) o compuestas, llamadas también expedientes (con características heterogéneas); seguidamente, atraviesan por un proceso de organización, transferencia y, finalmente, se define si estos son conservados, seleccionados o eliminados, de acuerdo con la información que poseen y su valor para la entidad.

Es así como se contribuye a garantizar la transparencia en la administración de la entidad, al igual que el seguimiento en la toma de decisiones, ya que las actividades de este proceso permiten tener disponibles los documentos que fueron producidos o recibidos en el marco de las funciones de la entidad.

7.1.2.1 Tipos de documentos

Existen diferentes tipos de documentos en las entidades; una de las primeras divisiones que pueden realizarse son documentos de archivo y de apoyo; los primeros son aquellos que se generan o reciben en razón de las funciones de una persona o entidad, y son objeto de aplicación de los procesos de gestión documental. Dentro de esta categoría se tienen, por consiguiente, documentos, de acuerdo con los valores que tienen para la entidad: misional, legal, administrativo, jurídico histórico o científico. Por otro lado, los documentos de apoyo son la guía o base que tiene como finalidad orientar la ejecución de las funciones.

Los documentos también pueden ser de carácter público, clasificado, reservado o confidencial, y esto dependerá de la información que contengan, la entidad que los genere y sus políticas de tratamiento de información. Así, en los procesos de gestión documental también se tienen documentos con información personal. Colombia se cuenta con leyes que protegen los datos personales, conservados o consignados en algún soporte documental, como la Ley General de Archivos 594 de 2000, mediante la cual se restringe el acceso a documentos de carácter reservado; la Ley 1268 de 2008, donde se dictan algunas disposiciones de habeas data y el acceso a datos financieros, crediticios, comerciales; la Ley 1581 de 2012, donde se dictan disposiciones generales sobre la protección de datos personales; la Ley 1712 de 2014, conocida como ley de transparencia, entre otras disposiciones.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

7.1.2.1 Principios de la gestión documental

De acuerdo con los principios establecidos en el Decreto 1080 de 2015 de Colombia, complementado con las disposiciones de la NTC-ISO 15489 que versa sobre *Información y documentación; gestión documental*, los principios que enmarcan la gestión documental son:

Figura 3 Principios de la gestión documental



Fuente: Elaboración propia

- *Eficiencia y economía*: para la optimización de recursos y lograr que estos sean eficientes, se plantea que deben ser generados sólo aquellos documentos que sean necesarios para la gestión de la entidad y que se produzcan en el desarrollo de alguna función; de esta forma, se contribuye a la economía, ya que lo que se pretende es que las entidades evalúen sus costos derivados de la gestión de documentos y puedan definir estrategias para el ahorro de los recursos implementados en el proceso.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- *Oportunidad*: se deberá garantizar la disponibilidad de los documentos, de manera que cuando se requieran, estos puedan ser consultados y así contribuir con el fin de los archivos: facilitar el acceso a la información. Las entidades implementarán los servicios de información necesarios para disponer los documentos a la comunidad en general, usuarios internos y externos de acuerdo con las necesidades de cada uno.
- *Transparencia*: por medio de una correcta gestión documental, la entidad podrá hacer sus trámites de manera transparente, pues los documentos sirven como evidencia si cumplen las características de autenticidad, es decir, «que el documento es lo que afirma ser, que ha sido creado o enviado por la persona que afirma haberlo creado o enviado, y que ha sido creado o enviado en el momento que se afirma» (Archivo General de la Nación, 2012); para ello, cada entidad deberá regular y controlar la producción, recepción, gestión y mantenimiento de los documentos, de manera que estos sean creados por la persona idónea, al momento de ejecutar alguna función o actividad, sin que se genere inalterabilidad, de suerte al cumplir con las cualidades enunciadas, este pueda ser consultado sin ningún inconveniente y se garantice la transparencia, de cara al usuario (persona natural o jurídica).
- *Protección del medio ambiente*: evitar la producción de documentos que no son necesarios para la entidad y definir el soporte en que será producido el documento, de acuerdo con sus tiempos de retención, son algunas de las actividades que pueden realizarse desde la gestión documental, para contribuir con la protección del medio ambiente; además, de actividades como el reciclaje y la optimización de recursos. Como ejemplo de esto, en Colombia se implementan las estrategias y políticas de cero papel, las cuales fueron dispuestas con el fin de proteger el medio ambiente, según la Directiva Presidencial 04 de 2012 con el asunto de *eficiencia administrativa y lineamientos de la política cero papel en la administración pública*. En esta línea, las impresiones de los documentos se realizarán cuando sea estrictamente necesario, situación que puede preverse desde la planeación documental; además, se pueden ejecutar actividades relacionadas con la correcta disposición de los elementos o desechos del archivo, como ganchos metálicos y papel reciclable, así como la optimización de recursos, tanto de papelería como de recursos tecnológicos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- *Cultura archivística*: este principio busca generar y promover el reconocimiento del archivo y de los procesos de gestión documental realizados en la entidad, para que las personas alcancen mayor nivel de conciencia, respecto a la importancia de estos; para ello, es necesario realizar procesos de sensibilización y capacitación al personal de la entidad.
- *Modernización*: en el proceso de gestión documental se hará uso de las tecnologías de la información y las comunicaciones; deberán permitir gestionar los documentos desde su creación, de acuerdo con una serie de responsables de su generación y gestión, establecidos previamente; asimismo, deben estar enfocados en las necesidades y ser afines con los objetivos de la entidad, los usuarios y demás partes interesadas; facilitar la toma de decisiones basadas en evidencias, permitir la producción de documentos, su gestión, uso y almacenamiento. Sumado a lo anterior, deben garantizar la interacción con las personas pertenecientes a la entidad, de acuerdo con ciertos niveles de acceso, previamente definidos y, de esta manera, garantizar la protección de los documentos y los datos, teniendo en cuenta los tipos que se gestionarán y almacenarán en los sistemas.
- *Vínculo archivístico*: la generación, y posterior agrupación de los documentos, debe corresponder a un mismo trámite, es decir, a la hora de conformar las unidades documentales no se deben combinar los trámites que dieron origen a los mismos, porque así es como se garantiza que el vínculo que poseen los documentos, ya sea por ser parte de un mismo trámite, de un fondo o serie documental, sea garantizado, con lo cual se pretende, además, que se genere una correcta integración de los documentos pertenecientes a un mismo contexto de producción.
- *Control y seguimiento*: como en todo proceso, es necesario mantener un control y una evaluación constante de la gestión documental en la entidad, que permita hacer seguimiento al proceso y se den las acciones de mejora a tiempo para ser tenidas en cuenta. Lo dicho, viabilizará el cumplimiento a cabalidad de los requisitos normativos y tecnológicos, a la hora de gestionar, almacenar, conservar y disponer los documentos.
- *Interoperabilidad*: el sistema de gestión documental deberá ser interoperable, esto es, estar integrado a los demás sistemas con los que cuenta la entidad —de calidad, de gestión

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

humana, de riesgos—, así como los diferentes sistemas donde se gestionen o tramiten documentos; esto con el fin de garantizar no solo una mayor fluidez, sino, también, una mayor complementariedad, teniendo en cuenta que al integrar los sistemas se estaría obteniendo una información integral y completa.

- *Orientación al ciudadano*: teniendo en cuenta la finalidad última de los archivos, a saber, servir como fuente de acceso a la información, la gestión documental debe prestar un servicio para el ciudadano. Así que el hecho de caracterizar a los usuarios, cobra gran relevancia, pues, de esta forma, se garantiza la atención especializada, de acuerdo con las características propias de cada tipo de usuario, para que estos puedan solventar los requerimientos en cuanto a acceso a los documentos.

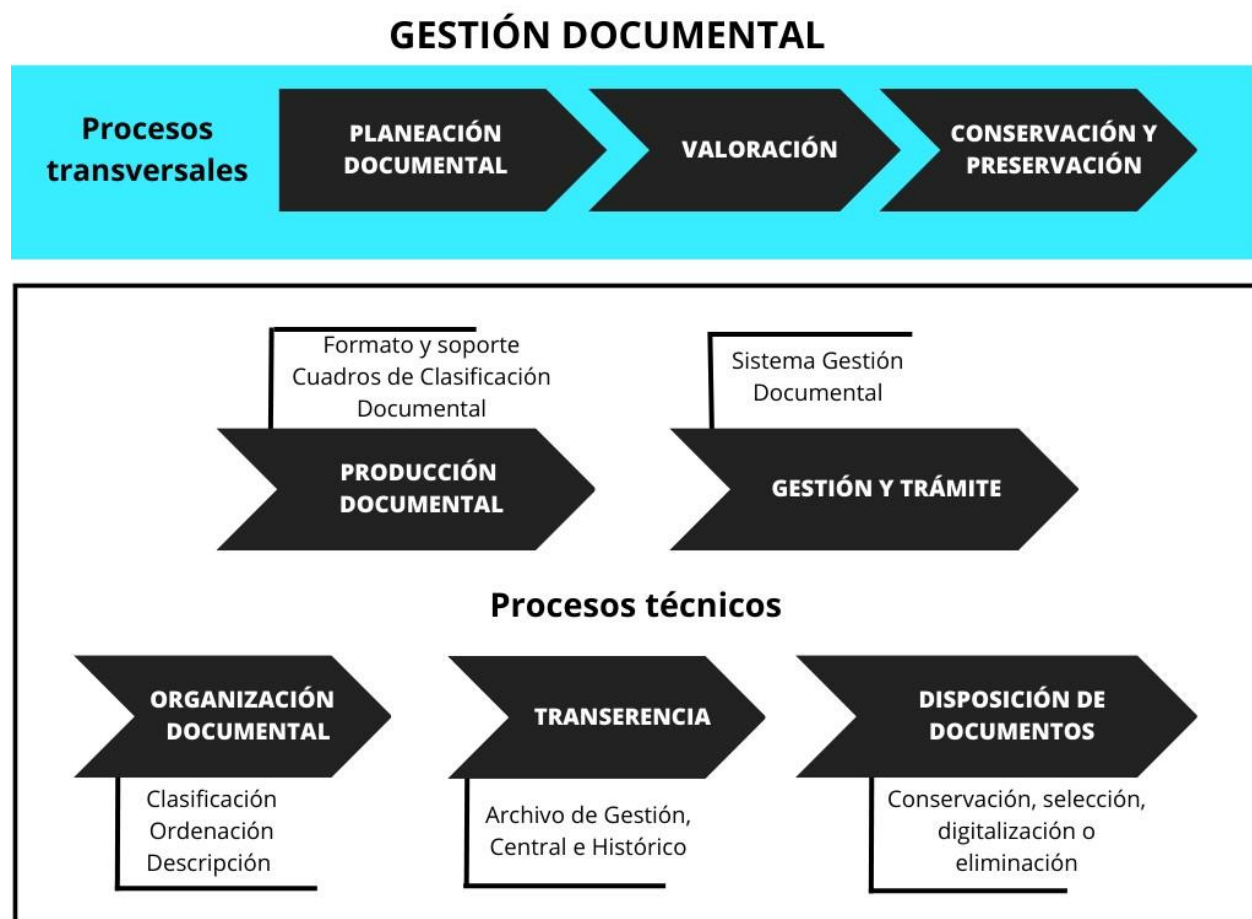
7.1.2.2 Procesos de la gestión documental

Para describir los procesos de la gestión documental se tuvo en cuenta lo expuesto en la NTC- ISO 15489, el Decreto 1080 de 2015 de Colombia, el Manual para la Implementación de un Programa de Gestión Documental y el Modelo de Gestión Documental y Administración de Archivos, todos ellos definidos por el Archivo General de la Nación, en virtud a que en estos recursos de información se encuentra el detalle, el paso a paso (procesos) de la gestión documental.

En primer lugar, se ilustran los procesos, tanto transversales como secuenciales; posteriormente, se desarrolla cada uno, teniendo en cuenta que estos son una guía para la gestión documental (figura 4):

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Figura 4 Procesos gestión documental



Fuente: Elaboración propia

- *Procesos transversales de la gestión documental:* considerados como estratégicos para el desarrollo de esta investigación, brindan orientaciones para el desarrollo de los demás procesos secuenciales:

A. Planeación documental: contempla aspectos previos a la producción, relacionados con la necesidad de crear el documento, bajo qué función se está creando, si este se encuentra registrado en el instrumento archivístico; así mismo, tablas de retención documental (TRD), que se definen como el «listado de series, con sus correspondientes tipos documentales, a las cuales se asigna el tiempo de permanencia en cada etapa del ciclo vital de los documentos» (Archivo General de la Nación, 2006), así como su conservación; además, se plantea la asignación de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

metadatos de contenido, estructura y contexto para los documentos de archivo. En este sentido, un documento debe ser originado o recibido por la entidad, en razón de sus funciones; de lo contrario, sería solo un documento de apoyo o, incluso, ser ajeno a la entidad, y estos no serían de gran utilidad a la hora de la toma de decisiones.

Ahora bien, este proceso se apoya en los instrumentos archivísticos, no solo las TRD; en este proceso también cobran relevancia las tablas de valoración documental (TVD), que se definen como el «listado de asuntos o series documentales a los cuales se asigna un tiempo de permanencia en el archivo central, así como una disposición final» (Archivo General de la Nación, 2006) y los cuadros de clasificación documental (CCD), entendidos como un «esquema que refleja la jerarquización dada a la documentación producida por una institución y en el que se registran las secciones y subsecciones y las series y subseries documentales» (Archivo General de la Nación, 2006), con el fin de realizar la planeación técnica de la gestión documental y, de esta forma, tener un panorama y un control general de la producción, valoración y disposición final de los documentos. En este sentido, las actividades que se contemplan como parte de este proceso, van desde la definición de formatos, identificación de metadatos, firmas autorizadas, elaboración de flujos documentales, hasta la automatización de procesos.

- B. Valoración documental:** en este proceso se determinan los valores de los documentos, registrados en las TRD o TVD, teniendo en cuenta los valores asignados para cada una de las series documentales identificadas, partiendo del contexto, la misión y diferentes actividades que ejecuta la entidad que produjo o recibió los documentos que hacen parte de estos instrumentos. Los valores que pueden asignarse son los primarios, donde se contempla el carácter administrativo, legal y técnico de los documentos y los secundarios, donde se plantea un valor histórico, científico y cultural, y, de acuerdo con esto, se define la disposición final de los documentos. La ejecución de este proceso deberá estar a cargo de un grupo interdisciplinario, que tenga la presencia y experticia, mínimamente, de un archivista, un historiador y demás profesionales que se requieran, de acuerdo con las funciones que ejerza la entidad, bajo las que se producen los documentos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

C. *Conservación y preservación documental*: este proceso tiene como finalidad velar por la conservación de los documentos, unidades documentales, independiente del soporte en que se encuentren almacenados. El almacenamiento de documentos debe ser realizado de manera que se evite la pérdida, destrucción, acceso no autorizado, entre otros riesgos a los que se ven expuestos los documentos, tanto físicos como electrónicos; además, en el lugar donde se encuentre dicho soporte, debe realizarse seguimiento a las condiciones de ambiente, el uso de herramientas de protección, seguimiento a la rutina de seguridad de los documentos, entre otros aspectos relevantes para garantizar la conservación de estos.

Parte de este proceso radica en el uso y almacenamiento de los documentos con fines de conservación; de ahí que se planteen como alternativas, la creación de copias, incluso en otros soportes; la migración o conversión, el proceso que deberá quedar documentado e informado a las partes interesadas; para su logro, se deberá poner mayor atención y cuidar de no eliminar los documentos ni sus metadatos en un sistema o formato, hasta que en su lugar de destino se efectúe, a cabalidad, la migración o conversión. El instrumento archivístico definido para este proceso es el Sistema Integrado de Conservación (SIC), conformado por el Plan de Conservación y el Plan de Preservación Digital, donde se determinan las actividades a realizar, los responsables y los recursos necesarios para ejecutarlas, así como el cronograma.

- *Procesos técnicos de la gestión documental*: a continuación, se desarrollan los procesos secuenciales de la gestión documental:

A. *Producción documental*: en este proceso se contempla la parte más técnica de generar un documento, su formato, la impresión de este o el sistema en el que será creado y se traza una línea de ruta para la conservación de este, es decir, si este será generado en papel, electrónico o qué soporte es el más adecuado de acuerdo con el tipo de contenido que se planea generar y gestionar, además de la identificación de los responsables. La generación de un documento deberá responder a la ejecución propia de una función en la entidad, y este deberá estar acompañado de ciertos metadatos que permitan identificar o documentar su creación, esta se encontrará

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

normalizada por los Cuadros de Clasificación Documental, de acuerdo con registros maestros de documentos y demás directrices impartidas por el área y el sistema integrado de gestión.

- B. Gestión y trámite:** se hace referencia a capturar el documento en un sistema, cuando la valoración de este lo indique, para gestionarlo dentro del mismo, por motivos de agilidad y conservación, ya que los sistemas permiten el acceso oportuno a los documentos, así como mantenerlo almacenado y accesible. Una vez generado, recibido o capturado el documento, comienza su distribución de acuerdo a su finalidad, para ello es necesario tener claros los responsables y aquellas personas que intervendrán en el proceso; durante esta etapa se recomienda tener definidos los niveles de acceso a la información, para ello, otro de los instrumentos archivísticos reglamentados en el Decreto 1080 de 2015, es la Tabla de Control de Acceso; además, se recomienda tener clara la trazabilidad del documento para gestionarlo en el sistema.

Este proceso abarca actividades como la distribución de los documentos, el trámite, su acceso y consulta, hasta la finalización del trámite para el que fue creado. Se cuenta entonces con gestores documentales para la recepción, distribución y resolución de comunicaciones y con Sistemas de Gestión de Documentos, que permite hacer seguimiento al trámite del documento de una manera más integral, ya que este permite mayor control desde la creación del documento, no solo comunicaciones, hasta la culminación del trámite.

- C. Organización documental:** para la organización es necesario hacer vínculos entre documentos ya sea de manera individual o agrupaciones. Para generar estos vínculos es necesario partir de la Clasificación Documental y, por ende, del instrumento archivístico Cuadros de Clasificación Documental, que permiten identificar bajo qué trámite y dependencia fueron creados y así realizar correctamente las agrupaciones documentales que dan origen a las unidades documentales y expedientes. Una vez realizada la *clasificación*, se continúa con la *ordenación*, que consiste en tomar cada unidad documental conformada y dar un

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

orden a cada documento, respetando la secuencia del trámite y la producción documental, por lo general se da un orden cronológico. Finalmente, la *descripción* permite, principalmente, la recuperación de documentos. En este punto, lo más usado es el Formato Único de Inventario Documental, establecido por el Archivo General de la Nación; sin embargo, existen normas de descripción internacionales como la Norma Internacional General de Descripción Archivística [ISAD(G)], que pueden ser usadas, no solo para la recuperación de documentos, sino también de imágenes y demás recursos de información con que cuente el archivo.

- D. *Transferencia:*** Los documentos pasan por unas fases, estos son generados y tramitados en los archivos de gestión (oficinas), en esta fase se generan expedientes y unidades documentales que de acuerdo a unas características, evaluadas en el proceso de valoración, se asignan tiempos de retención, de manera que, una vez cerrado el expediente o unidad documental, este inicia el conteo de su tiempo de conservación en el archivo de gestión y una vez cumplido pasa al archivo central; en esta fase solo es almacenado y consultado, para luego pasar al proceso de eliminación, selección o conservación total, de esta manera solo pasan a la última fase aquellas unidades documentales que de acuerdo a sus valores para la entidad, la ciencia o la cultura, serán conservados en el archivo histórico, todo esto de acuerdo con lo establecido en las TRD o TVD.

Para la transferencia de documentos, es necesario establecer un cronograma, para que este proceso se dé ordenadamente en la entidad. Esta planeación facilitará proyectar el espacio y recursos necesarios para la custodia de los documentos. Ahora bien, en el caso de los documentos electrónicos, de acuerdo con lo establecido por el Archivo General de la Nación en el Modelo de requisitos para la implementación de un Sistema de Gestión de Documentos Electrónicos de Archivo (SGDEA), la entidad deberá contar con un sistema, que deba estar parametrizado para dar cumplimiento a los diferentes procesos de la gestión documental, de esta manera se podrán configurar los tiempos de retención y la disposición final de los documentos, teniendo en cuenta el requisito con el que se debe cumplir: facilitar la

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

parametrización de las Tablas de Retención y Valoración Documental de manera que la conservación, eliminación o selección se haga de forma automática a través de metadatos.

- E. Disposición de documentos:** este proceso es ejecutado en la fase de archivo central, de acuerdo con las directrices establecidas en las TRD y TVD, donde se determinan aquellos documentos que son de conservación permanente (pasan al archivo histórico), o se deben seleccionar para guardar una muestra, digitalizar o eliminar, de acuerdo con los valores de los documentos determinados durante el proceso de valoración documental. Cabe resaltar que los métodos de selección o muestreo de los documentos serán definidos en los instrumentos mencionados, así como el porcentaje de esta selección. Valga advertir que antes de realizar alguna eliminación, es necesario publicar un inventario donde la entidad informe a la comunidad en general los documentos que están próximos a eliminarse, de manera que se garantice la transparencia en su gestión y se evite la destrucción de información que pueda ser relevante para alguna área o persona particular, asimismo, se deben definir los métodos de destrucción de los documentos, verificando la normatividad que avale este proceso y se debe dejar un acta que sustente el proceso realizado.

7.2 Relaciones entre la gestión de datos y la gestión documental

Partiendo de las características expuestas de la gestión de datos y la gestión documental, en este apartado se elabora un relacionamiento entre sus definiciones, tipos, principios y procesos, con el fin de identificar las similitudes, diferencias y complementariedades que existen entre estos; además, se realiza un análisis de cada elemento con la generalidad de la gestión de datos o la gestión documental, según corresponda.

7.2.1 Gestión de datos y gestión documental: definiciones

Tanto la gestión de datos como la gestión documental se refieren a la aplicación de una serie de estrategias, normas, políticas y procesos, aplicadas al ciclo de vida de los datos y documentos; ambos ciclos parten desde la generación, recepción o extracción, pasando por procesos de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

organización, clasificación, agrupación, visualización y uso, hasta determinar si los datos o documentos serán conservados y almacenados y, solo para el caso de los documentos, si estos serán eliminados, de acuerdo con las fuentes usadas en el desarrollo de esta investigación.

Cabe resaltar que la terminología corresponde a las necesidades propias de cada proceso; no obstante, se identifica una alta similitud en las definiciones de gestión de datos y gestión documental, dado que contemplan actividades afines, como los procesos de generación o producción, agrupación y almacenamiento, enfocados en cada tipo de dato y documento, de acuerdo con su interés particular. A continuación, se ilustra la relación entre las definiciones, en la cual se evidencia la afinidad entre estos, lo que permite dar coherencia y posibilidades al planteamiento de las estrategias y, en general, al desarrollo de esta investigación:

Figura 5 Gestión de datos y gestión documental: definiciones



Fuente: Elaboración propia

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

7.2.2 Tipos de datos y documentos

Existen diferentes tipos de datos y de documentos, con necesidades específicas a la hora de tratarlos o gestionarlos; por tanto, tener presente estas singularidades es importante, no solo porque marcan un punto relevante de diferenciación y relacionamiento entre la gestión de datos y la gestión documental, sino también para garantizar su adecuado tratamiento desde la perspectiva archivística. En la tabla 1 se enuncian las categorías dentro de los tipos de datos y documentos, y si estas pueden encontrarse dentro de la gestión de datos y la gestión documental:

Tabla 4: Tipos de datos y documentos

Tipos de datos	Gestión de datos	Gestión documental
Simple	✓	
Compuestos	✓	
Maestros	✓	
Transaccionales	✓	✓
Estructurados	✓	
No estructurados	✓	✓
Semiestructurados	✓	✓
Tipos de documentos	Gestión de datos	Gestión documental
De archivo	✓	✓
De apoyo		✓
Tipos de datos y documentos de acuerdo con su nivel de acceso	Gestión de datos	Gestión documental
Público	✓	✓

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Clasificado	✓	✓
Reservado	✓	✓
Confidencial	✓	✓
Privados	✓	✓
Personales	✓	✓

Fuente: Elaboración propia

Como coincidencias o similitudes, se refleja que los datos transaccionales, aquellos que se gestionan en programas donde son creados, tramitados e, incluso, almacenados, se encuentran presentes dentro de la gestión documental, más específicamente para los documentos electrónicos, como es el caso de los datos resultantes de una transacción bancaria, en la que es posible evidenciar fecha, monto, número de cuenta, entre otros; estos tipos de datos, de acuerdo con el ejemplo enunciado, podrían consolidarse en un documento, como un comprobante de egreso o ingreso electrónico.

Los datos no estructurados —que hacen referencia a documentos, textos, imágenes y audiovisuales— y los datos semiestructurados —como la información de correos electrónicos— también se encuentran presentes en la gestión documental, siendo esto más común en el documento electrónico, ya que en este se encuentran subsumidas las fotografías, los correos electrónicos; es más, materializarse notoriamente en los archivos especializados en los que se encuentran diferentes formatos, aun audiovisuales, objetos de la gestión documental.

Se agrega a lo anterior, otras coincidencias se encuentran en los niveles de acceso, pues, en ambos casos, se plantea el tratamiento de datos o documentos de carácter público, privado y reservado; para esto, se cuenta con el instrumento de control de acceso, la parametrización de un sistema que permita categorizar el ingreso a ciertos datos o documentos, entre otras herramientas, definidas para el caso de Colombia de acuerdo con la normatividad sobre protección de datos, como la Ley 1266 de 2008 y la Ley 1712 de 2014.

Por otro lado, se encontró que los tipos de documentos enunciados como de apoyo y de archivo no se encuentran explícitos en la gestión de datos; empero, de acuerdo con la definición de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

cada una de estas categorías y haciendo referencia a los documentos que sirven de ayuda para la gestión de la entidad, sumados a aquellos que surgen en el marco de la ejecución de las funciones de esta —sin perder de vista que los datos son producidos o extraídos en el marco de alguna función, actividad o proceso de la entidad— es preciso afirmar que estos son datos de archivo. Bien podrían servir de apoyo, toda vez que los datos son usables y reutilizables en alguna gestión específica.

Paralelamente, como diferencias entre los tipos de datos y documentos, se encuentran aquellos que son propios de alguno de ellos, por ejemplo, los datos simples y compuestos. Los mismos pueden encontrarse inmersos en bases de datos, ya que estos corresponden a números enteros, reales, aun listados para el caso de los datos compuestos. En esta línea no se logró mostrar ningún tipo de relación con la gestión documental. Además, se tienen los datos maestros con los que se busca dar orientación y contexto a los otros datos extraídos o producidos, y los estructurados, consolidados de forma normalizada en bases de datos, que no son objeto de la gestión documental, según la teoría y normativa, analizada. Pese a todo, es relevante desarrollar estrategias para gestionarlos, ya que actualmente es muy común encontrar en las entidades, bases de datos estructuradas con información relevante, es más, misional, que, por ende, deben gestionarse de modo integral.

7.2.3 Gestión de datos y gestión documental: principios

Los principios son la base sobre la cual se establecen los procesos de la gestión de datos y la gestión documental. El curso de esta búsqueda, es necesario identificar los puntos de encuentro y las diferencias entre estos para fundamentar aún más las relaciones que tienen. A continuación, se enuncian los principios, tanto de la gestión de datos como de la gestión documental y se indica, en primera instancia, si existe relación entre ellos:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Tabla 5: Relación entre los principios

Principios	Eficiencia y economía	Oportunidad	Transparencia	Protección del medio ambiente	Cultura archivística	Modernización	Vínculo archivístico	Control y seguimiento	Interoperabilidad	Orientación al ciudadano
Los datos son un activo con propiedades únicas		✓								
La gestión de datos es la gestión del ciclo de vida										
Gestionar datos significa gestionar la calidad de los datos			✓					✓		
Se necesitan metadatos para administrar los datos										
La gestión de datos es transversal e interdisciplinar					✓					
La gestión de datos requiere una perspectiva integral					✓					
La gestión de datos incluye su gestión de riesgos										
Los requisitos de gestión de datos deben impulsar las decisiones de tecnología de la información						✓				

Fuente: Elaboración propia.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Para los Principios de la gestión de datos y la gestión documental, se realizó un comparativo entre ambos, dando como resultado lo expuesto en la tabla 2, donde se puede observar que cinco de los ocho principios de la gestión de datos se relacionan con los principios de la gestión documental. En este contexto, cabe recordar que en el principio *los datos son un activo con propiedades únicas*, donde queda establecido que los datos toman valor cuando se agrupan, y se extrae información para diferentes fines, tiene relación con la gestión documental, más específicamente con el principio de *oportunidad*, el cual hace mención a la disponibilidad de los documentos y la relevancia que tienen estos para un usuario o quien desee consultarlo; es decir, extraer información del mismo, ya que este es el momento en el cual tanto datos como documentos cobran valor.

Una coincidencia que suma a las anteriores está en el principio *gestionar datos significa gestionar la calidad de los datos*, cuyo propósito está puesto en que los datos sean usables, fiables y completos. Este principio está relacionado con los principios *de transparencia, control y seguimiento* de la gestión documental, ya que con estos se pretende que los documentos sirvan como evidencia, garantizando que estos cumplan las características de autenticidad; en otros términos, que este no haya sido alterado, así como la necesidad de mantener una evaluación constante de la gestión documental en la entidad, con el fin de realizar un seguimiento al proceso y se puedan disponer las acciones de mejora que apunten a garantizar la calidad de los documentos. Pese a que este relacionamiento no es explícito, esto se pudo corroborar, posterior al análisis de los principios enunciados.

Uno más, el principio *la gestión de datos es transversal e interdisciplinar*, por tanto, involucra todas las áreas de la entidad. Se busca contar con una amplia gama de profesionales en el proceso; adicionalmente, está relacionado con el principio *de cultura archivística*, de la gestión documental, que se ocupa de que todas las personas de la entidad reconozcan el valor de los documentos, del archivo, y se trabaje en comunidad para lograr una correcta gestión de documentos; así mismo, el principio *de la gestión de datos requiere una perspectiva integral*, es decir, que se implemente cada proceso de la gestión de datos en las entidades.

El último de los principios de la gestión de datos se refiere a que *los requisitos de gestión de datos deben impulsar las decisiones de tecnología de la información*, en el cual se plantea un papel importante de las tecnologías como herramientas necesarias para el tratamiento de datos, a

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

lo largo de su ciclo de vida, y el fortalecimiento de la gestión de datos, de manera que estas se conciban como una oportunidad. Cabe señalar que se vincula con el principio *de modernización de la gestión documental*, ya que este también reconoce la importancia de dichas herramientas y los sistemas para la gestión y protección de los documentos electrónicos y los datos; lo anterior, está enfocado en las necesidades de las entidades y los usuarios; en ambos casos es evidente el uso de tecnologías para la gestión de datos y la gestión documental.

Llegados hasta aquí, es necesario aclarar que esto es solo uno de los análisis que puede extraerse de los principios, ya que también es útil enunciar si dentro de la gestión de datos y gestión documental se encuentran comprometidos, de alguna forma, los principios expuestos, es decir, un relacionamiento no solo entre los principios sino algo más amplio, de manera que se tengan en cuenta los demás ítems —como los conceptos, procesos y tipos—:

Tabla 6: Principios en la gestión de datos y la gestión documental

Principios de la gestión de datos	Gestión de datos	Gestión documental
Los datos son un activo con propiedades únicas		
La gestión de datos es la gestión del ciclo de vida	✓	✓
Gestionar datos significa gestionar la calidad de los datos		
Se necesitan metadatos para administrar los datos	✓	✓
La gestión de datos es transversal e interdisciplinar		
La gestión de datos requiere una perspectiva integral		
La gestión de datos incluye su gestión de riesgos	✓	✓
Los requisitos de gestión de datos deben impulsar las decisiones de tecnología de la información		

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Principios de la gestión documental	Gestión de datos	Gestión documental
Eficiencia y economía	✓	✓
Oportunidad		
Transparencia		
Protección del medio ambiente	✓	✓
Cultura archivística		
Modernización		
Vínculo archivístico		✓
Control y seguimiento		
Interoperabilidad	✓	✓
Orientación al ciudadano		✓

Fuente: Elaboración propia

En esta misma óptica, el principio *la gestión de datos es la gestión del ciclo de vida*, si bien no tiene relación particular con ningún principio de la gestión documental, este se encuentra inmerso dentro de esta; al igual que los documentos, los datos también son gestionados durante todo su ciclo de vida, pese a que estos ciclos son diferentes, acorde con lo expuesto en el apartado de las definiciones.

De la misma manera ocurre para el principio *se necesitan metadatos para administrar los datos*, ya que los metadatos se usan para identificar las características de los documentos, brindar información sobre su contenido y estructura, así como identificar un contexto con el fin de gestionarlo y conservarlo a lo largo del tiempo; situación presente en la gestión documental desde el proceso de planeación, en el cual se plantea la asignación de metadatos de contenido, estructura y contexto para los documentos de archivo; en el proceso de conservación, relacionado con la conservación del documento, así como los metadatos que los integran y en el proceso de producción, el cual se menciona que debe estar acompañado de metadatos que permitan identificar o documentar la creación de los documentos. Así las cosas, es posible afirmar que se encuentra un

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

punto de relación de un principio de la gestión de datos con estos procesos de la gestión documental.

Así mismo con el principio *la gestión de datos incluye su gestión de riesgos*, si bien es propio de la gestión de datos, también se encuentra presente en la gestión documental, pues a lo largo de la planeación y formulación de los procesos de la gestión documental, se busca identificar los riesgos a los que están expuestos los documentos, de forma que se mitiguen y se garantice la continuidad de estos.

Para el caso de los principios de la gestión documental en relación con la gestión de datos, en general, se logró constatar que la *eficiencia y economía*, y la *protección del medio ambiente* son principios con los cuales se busca la optimización de los recursos, así como la producción consciente y medida de los documentos, aspectos que son parte de la gestión de datos, al afirmarse que su producción se realiza teniendo en cuenta las funciones que se ejecutan al interior de las entidades y que esta se hace en sistemas, elementos de la gestión de datos que permiten que se tenga una relación con los principios enunciados de la gestión documental. Además, el *principio de interoperabilidad* se relaciona con la gestión de datos, más específicamente con el proceso de integración de datos e interoperabilidad, ya que en ambos casos se contempla la implementación e integración de los sistemas, es decir, que estos sistemas pueden relacionarse, comunicarse entre sí o ejecutar tareas en conjunto.

Para concluir este apartado, se identifican los principios de la gestión documental que no se encuentran presentes en ningún elemento de la gestión de datos, pero son complementarios; en esta dirección, se afirma que el *vínculo archivístico* es fundamental para lograr una correcta agrupación de datos, conforme a las estructuras orgánicas y funcionales de las entidades y que el *principio de orientación al ciudadano*, es necesario tenerlo en cuenta para la gestión de datos, toda vez que estos, en su proceso de explotación, son el insumo para un servicio de información en la entidades, donde se cuenta con usuarios internos o externos, que además requieren ser caracterizados, con el fin de atender y responder sus requerimientos, eficazmente.

7.2.4 Gestión de datos y gestión documental: procesos

En este apartado, se desarrollan las relaciones entre los procesos, tanto estratégicos como técnicos, de la gestión de datos y la gestión documental. En primer lugar, se analizaron los procesos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

estratégicos y transversales, que contribuyen con el direccionamiento de los demás procesos técnicos, ya que estos buscan definir los responsables de los datos y documentos durante todo el ciclo de vida, según el caso. Así mismo, estos procesos se ejecutan con el fin de garantizar datos y documentos de calidad, disponibles, fiables y auténticos.

7.2.4.1 Gestión de datos y gestión documental: Procesos estratégicos y transversales

A continuación, se muestran las relaciones encontradas entre los procesos estratégicos y transversales:

Tabla 7: Gestión de datos y gestión documental: procesos estratégicos

Procesos de la gestión Documental	Gestión de datos	Gestión documental
Planeación	✓	✓
Valoración		✓
Conservación y preservación		✓
Procesos de la gestión de datos	Gestión de datos	Gestión documental
Gobernanza	✓	✓
Calidad	✓	✓

Fuente: Elaboración propia

De acuerdo con lo expuesto en la tabla 4, se logra identificar que de los procesos estratégicos enunciados, tanto de la gestión de datos como de la gestión documental, se tiene relación solo entre tres de ellos, el proceso de *planeación* y de *gobernanza*, ya que para ambos casos se plantea la definición de responsables y la identificación de datos o documentos, según aplique, que serán producidos en las diferentes áreas de las entidades y, por ende, son relevantes para las mismas; estos procesos se materializan en instrumentos como la matriz de datos críticos, para el caso de la

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

gestión de datos, donde se relacionan o definen aquellos datos que son producidos o recopilados por las entidades, y los cuadros de clasificación documental, para el caso de la gestión documental, en donde se relacionan todos los documentos, productos de la gestión diaria de las entidades. De lo anterior, se confirma la relación entre ambos procesos, incluso, en los instrumentos en las que estos se materializan.

La *calidad* como proceso estratégico sólo se encuentra dentro de la gestión de datos; sin embargo, esta tiene relación con la gestión documental, más específicamente, en los principios de transparencia y control y seguimiento, ya que en estos se mencionan los documentos con características de autenticidad, obtenidos por el control sobre la producción, recepción, gestión y mantenimiento de los documentos, así como por la evaluación constante de la gestión documental, con acciones de mejora que permiten obtener documentos de calidad.

Es importante resaltar, no solo las diferencias, sino también las oportunidades, es decir, aquellos procesos transversales que no se encuentran relacionados con la gestión de datos, pero que se consideran relevantes para el tratamiento archivístico de los datos, de manera que se complemente la gestión de datos con los procesos transversales de la gestión documental, como es el caso de la *preservación* y la *valoración*, para lo cual se podrán desarrollar estrategias para su almacenamiento, acceso, condiciones de reutilización, y demás, que permitan y garanticen la preservación a largo plazo de los datos.

Referido al proceso de valoración, se considera como un elemento necesario para determinar la producción y disposición final de estos; esto es, definir aquellos que se deben generar y conservar y, así, evitar el almacenamiento de datos sin ninguna utilidad y con un gran volumen que solo satura los sistemas y repositorios donde se encuentran; igualmente, se podrá prevenir la eliminación de aquellos que sean de interés, no solo para las entidades, sino también para la comunidad en general.

7.2.4.2 Gestión de datos y gestión documental: procesos técnicos

Los procesos técnicos implican la ejecución de una serie de actividades para garantizar la gestión de datos y de documentos, abarcando cada fase de su ciclo de vida, que va desde la creación del dato o documento, hasta su uso y posterior eliminación o conservación permanente, al menos, para el caso de la gestión documental, ya que para los datos no se tiene definida una disposición final.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

A continuación, se muestran los procesos técnicos y su relación con la gestión de datos y la gestión documental (tabla 5):

Tabla 8: Gestión de datos y gestión documental: procesos técnicos

Procesos de la gestión de datos	Gestión de datos	Gestión documental
Arquitectura	✓	✓
Soporte de las operaciones	✓	✓
Protección de datos	✓	✓
Integración e interoperabilidad	✓	✓
Gestión de documentos y contenidos	✓	✓
Almacenamiento de datos e inteligencia comercial	✓	✓
Explotación de datos	✓	
Procesos de la gestión documental	Gestión de datos	Gestión documental
Producción	✓	✓
Gestión y trámite	✓	✓
Organización documental	✓	✓
Transferencia		✓
Disposición de documentos		✓

Fuente: Elaboración propia

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

En primer lugar, se hace mención de los procesos similares entre la gestión de datos y la gestión de documentos; en esta dirección y acorde con lo expuesto en la tabla 5, la *arquitectura* y la *producción* se encuentran en ambos casos, teniendo en cuenta que en dichos procesos se determina la generación de datos y documentos, y se definen los formatos a utilizar, los sistemas donde se gestionan, así como los requisitos y características con los que deben contar para que estos sean de valor para la entidad.

La *protección de los datos* también establece una relación, siendo este un proceso de la gestión de datos y parte del proceso de *gestión y trámite* de la gestión documental, en vista de en este último se contemplan aspectos relacionados con el nivel de acceso y normatividad que le aplique. En este sentido, es necesario tener pleno conocimiento de los responsables y aquellos que pueden intervenir en el proceso o acceder a los datos y documentos, logrando así una trazabilidad, plenamente identificada, así como la seguridad de estos.

La *integración e interoperabilidad* son otros de los procesos que están presentes en la gestión de datos y documental, puesto que ambos casos se usan sistemas que deben tener la capacidad de integrarse con otros, garantizar el intercambio de datos y documentos, así como permitir la ejecución de actividades de migración y almacenamiento. Lo anterior, para el caso de la gestión documental, se ve contemplado dentro del proceso de *gestión y trámite*, donde se hace referencia a la captura de documentos en un sistema para gestionarlos; para este caso, se cuenta con el desarrollo de sistemas de gestión de documentos electrónicos de archivo y un modelo de requisitos con el cual este sistema debe contar.

Para el proceso de *gestión de documentos y contenidos* de la gestión de datos, donde se contemplan actividades como *producción, gestión y trámite* y almacenamiento de documentos o datos no estructurados, se alcanza a evidenciar que pese a que se aspira abarcar varios procesos de la gestión documental, no logran ser lo suficientemente desarrollados, a diferencia de la gestión documental con diversidad de manuales, instrumentos y demás herramientas desarrolladas para el tratamiento integral de los documentos. Ahora bien, no se logró develar una relación o mención explícita donde se defina que la gestión documental tenga o contemple un proceso sobre gestión de datos; sin embargo, de acuerdo con el análisis realizado, se ha logrado determinar la presencia, en varios momentos, de principios, procesos y tipos de datos dentro de la gestión documental.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

El proceso de *almacenamiento e inteligencia comercial* tiene un punto de relación con el proceso de *organización* de la gestión documental, ya que en el proceso de la gestión de datos se plantean actividades de limpieza, donde se da un formato normalizado, preparando las bases de datos para llenar los campos vacíos o suprimir los datos duplicados; y en el proceso de organización documental, se hace referencia a agrupar o clasificar, ordenar y describir los documentos, conforme a unos parámetros previamente establecidos. Ambos procesos con la ejecución de actividades propias y ajustadas a las necesidades de los datos y de los documentos, apuntan a un mismo objetivo: que estos queden preparados para una fácil utilización por parte de los usuarios.

El proceso *soporte de las operaciones* de la gestión de datos está relacionado con el proceso de *gestión y trámite*, teniendo en cuenta que en ambos casos se contemplan las transacciones realizadas con los sistemas de administración de datos y documentos, partiendo de una gestión y almacenamiento en sistemas de información, actividad que debe ser monitoreada, debido al constante cambio tecnológico. Esto significa que se deberá garantizar el contexto. a los datos y documentos en los sistemas con que cuenten las entidades.

Una última similitud tiene ocurrencia con el proceso de *explotación de datos*, allí donde están contenidos el uso y la generación de valor de estos, con el fin de obtener beneficios para las entidades y usuarios, lo cual se vincula al proceso de *gestión y trámite* de la gestión documental, donde se contempla el acceso y consulta de los documentos, actividades con las cuales se saca provecho de estos. Esto implica que en ambos casos se debe contar con servicios de información, acordes con las necesidades y requerimientos de los usuarios, para que el acceso y uso sea eficiente.

En este contexto, las diferencias existentes radican en los procesos propios de la gestión documental, como es el caso del proceso de *transferencia*, tomando en consideración que para los datos no se tienen definidas fases o etapas por los que estos puedan atravesar a lo largo de su ciclo de vida, como el archivo de gestión, central e histórico, para el caso de los documentos, y para los datos sólo se concibe la generación, explotación y almacenamiento de estos. Otro de los procesos propios de la gestión documental es la *disposición de documentos*, porque, como se dijo, no se tiene definido un proceso que permita determinar si los datos son de conservación total o si estos deben ser eliminados, sin olvidar el valor que tienen para las entidades, motivo por el cual podrían conservarse un gran volumen de datos sin ningún uso o que sean irrelevantes —para las entidades, usuarios y comunidad en general— o, por el contrario, se podría incurrir en la eliminación de datos de gran valor.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Es importante reconocer la oportunidad que se tiene para complementar la gestión de datos con algunos procesos transversales y técnicos de la gestión documental, para citar, la *valoración*, *preservación* y *disposición final*, atendiendo las posibilidades y aplicabilidad que estos tienen; igualmente, podrían completarse los procesos de generación, almacenamiento y conservación de datos, conexos a las funciones de la entidad, buscando, entre otros, la optimización de recursos para la gestión de los mismos, que podrían darse con la implementación de estos procesos de la gestión documental en la gestión de datos, previa adaptación a las características y necesidades específicas de estos.

7.3 Estrategias para la gestión archivística de datos en Colombia

Una vez analizados los elementos contemplados para el desarrollo de esta investigación —definiciones, principios, tipos de datos y de documentos, y cada uno de los procesos que conforman la gestión de datos y la gestión documental, además del resultado del comparativo realizado, donde se identificaron diferencias, similitudes y complementariedades—, se presentan las estrategias para la gestión archivística de los datos en Colombia. Estas estrategias no son exhaustivas ni definitivas; por el contrario, son un primer acercamiento para el desarrollo paulatino y secuencial de este campo de conocimiento para la archivística en el país.

Para formular las estrategias para la gestión de datos desde la perspectiva archivística, se tuvo en cuenta el comparativo realizado, en un inicio se pretendía brindar estrategias para la generación, explotación, vinculación archivística, valoración, preservación y disposición final de los datos, teniendo en cuenta los grandes aportes y complementariedad que se puede hacer desde la gestión documental, teniendo en cuenta los puntos de conexión encontrados y las necesidades actuales de las entidades. Ahora bien, estrategias como la preservación, valoración y disposición final de los datos, fueron relevantes para esta investigación y para una gestión de datos desde la perspectiva archivística, finalmente se optó por abarcar un número específico de estrategias, con el fin de que estas fuesen formuladas y profundizadas en el desarrollo de una primera investigación y así dar paso a nuevos aportes y modificaciones a que haya lugar en una posterior investigación.

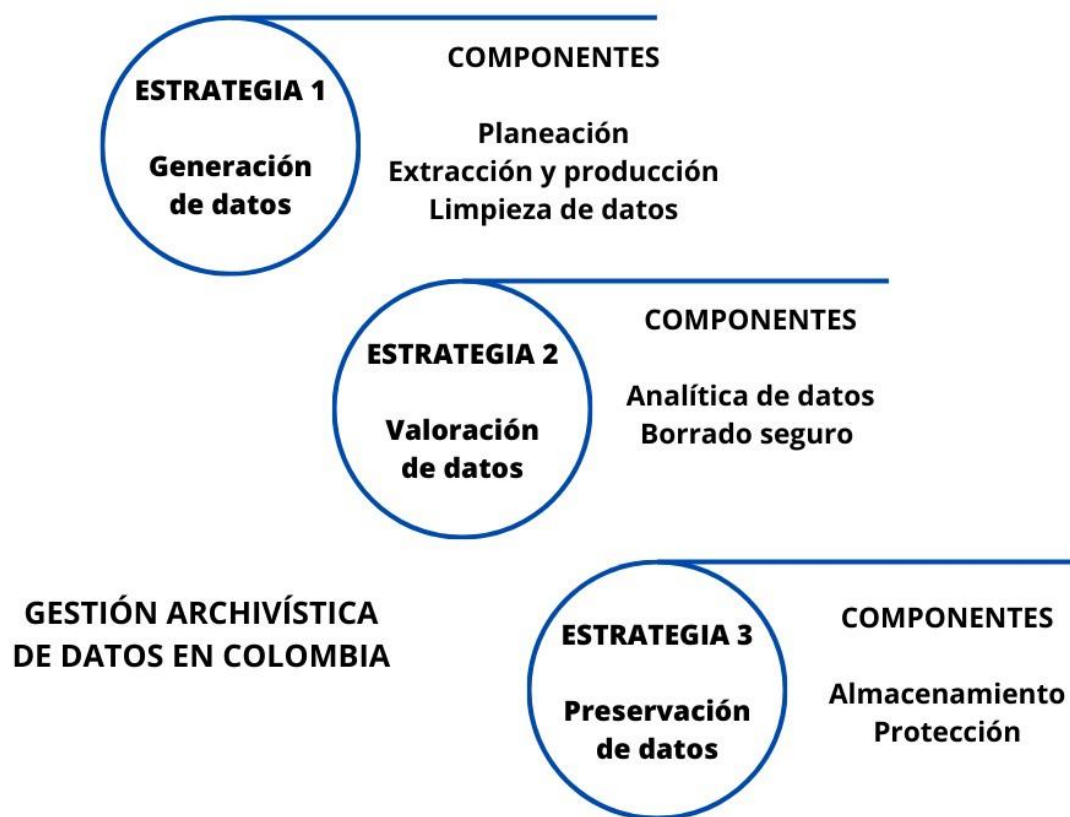
La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Ahora bien, estas fueron ajustadas, quedando más acordes a los hallazgos encontrados en el proceso de comparación, la experiencia y el conocimiento archivístico de base, así como el adquirido durante esta investigación sobre la gestión de datos. Como se logró observar se dejaron una serie de elementos relevantes, por fuera de las estrategias finales, sin embargo, como se mencionó estas son una primera propuesta, posteriormente será importante retomar temas como los mencionados por Jordi Serra sobre la explotación de datos como servicio: «el gran salto adelante en la gestión de la información se está dando en la forma como los usuarios la consumen e interactúan con ella» (Serra, 2021, p. 60).

Continuando el hilo de la exposición, en esta investigación se desarrollan una serie de estrategias para la gestión archivística de los datos en Colombia; en ellas se contempla el tratamiento de los datos durante todo el ciclo de vida, es decir, desde su generación, extracción o producción hasta su disposición final. Cada estrategia está conformada por una serie de componentes que posibilitan su comprensión y estos, a su vez, por actividades que permiten la ejecución de las estrategias. A continuación, se ilustran las estrategias:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Figura 6 Estrategias de la gestión archivística de datos en Colombia



Fuente: Elaboración propia

7.3.1 Generación de datos

Esta primera estrategia abarca el origen de los datos, momento en el cual se ejecuta su extracción y producción, en el marco de las funciones de las entidades, así como la garantía de su calidad. Siendo así, para el desarrollo de esta estrategia se plantean los componentes de planeación, extracción y producción, y, finalmente, su limpieza, en concordancia con los procesos de la gestión de datos ya desarrollados, de acuerdo con la guía DAMA-DMBOK2.

7.3.1.1 Componente planeación

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

La planeación es un componente previo a la extracción y producción de los datos, que consiste en verificar las necesidades de las entidades, áreas y responsables que dan origen a los mismos, con el fin de generar los datos necesarios para su gestión, así como definir un lenguaje normalizado para la extracción, producción y limpieza, mediante la implementación y elaboración de instrumentos como la matriz de metadatos y de datos maestros. Para el desarrollo de este componente es necesario ejecutar las siguientes actividades:

- A.** *Definir la matriz de datos maestros:* consiste en el diseño de una matriz de datos maestros, entendidos como aquellos cruciales para las entidades y las transacciones realizadas, esto es, aquellos datos resultantes de la ejecución de sus funciones. Esta matriz tiene como finalidad identificar los datos importantes que deben ser conservados y gestionados desde la perspectiva archivística.

- B.** *Elaborar la matriz de metadatos:* consiste en caracterizar los metadatos que serán extraídos o generados, así como la definición de su conformación o formato; para ello se tomarán metadatos de contexto, estructura y contenido. Esta actividad deberá estar en sintonía con la matriz de datos maestros; finalmente, son estos los que deberán contar con los metadatos definidos, de manera que se brinde información y se facilite su búsqueda, análisis y utilización.

7.3.1.2 Componente extracción y producción de datos

La extracción es el proceso mediante el cual los datos son tomados de sistemas informáticos que puede ser internos o externos a la entidad, para ser llevados a una base de datos, de acuerdo con la matriz de datos maestros y de metadatos. La producción de los datos es realizada de forma manual o por medio de formularios, encuestas y trámites, resultantes de la interacción con sistemas de gestión de datos. Para el desarrollo de este componente es necesario ejecutar las siguientes actividades:

- A.** *Capacitar a los productores de los datos:* consiste en dar a conocer los instrumentos de planeación: la matriz de datos maestros y la de metadatos. Esto permite que, en las entidades, los productores de datos tengan las orientaciones necesarias para realizar el

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

proceso de manera normalizada; así las cosas, con la ejecución constante de esta actividad se genera sensibilidad en los productores frente a la importancia que estos tienen para la gestión diaria de las entidades y para la toma de decisiones.

- B.** *Recolectar o producir los datos:* la recolección de los datos, la cual se realiza teniendo en cuenta la matriz de datos maestros y la de metadatos, para que los datos contemplados se estructuren en una base de datos; además, en este proceso de estructuración se deberán tener definidos, de un lado, los datos a tratar, su estructura, así como los responsables de esta actividad; de otro, tener identificadas las formas para tomarlos, llevarlos y agruparlos en la base de datos, es decir, contar con pleno conocimiento del sistema donde se encuentran aquellos datos que se requieren o definir la herramienta para recolectarlos. La producción de los datos, se realiza de forma manual o por intermedio de herramientas tecnológicas, donde no deben estar ausentes formas o formularios para la producción de datos. Cabe resaltar que los datos recolectados o producidos, llevados a una base de datos, responden a una temática y función específica; en esta línea, es importante mantener esta coherencia en su agrupación, siendo estos productos de las funciones de las entidades.
- C.** *Monitorear el cumplimiento e implementación de los instrumentos de la gestión archivística de datos:* esta actividad que antecede al componente de limpieza de los datos, está enfocada en dar cumplimiento a las directrices establecidas en la matriz de metadatos y en la matriz de datos maestros, por medio de un seguimiento constante, de forma que se generen o extraigan datos, acorde con estos instrumentos. Para ello se deberá hacer una revisión aleatoria y periódica de las bases de datos, mientras están siendo consolidadas, con el fin de verificar que se cuente con los registros tal y como fueron definidos en las matrices enunciadas.

7.3.1.3 Componente: Limpieza de los datos

La limpieza de los datos es el proceso mediante el cual se identifican y corrigen errores que pueden darse en el componente de extracción y producción; quiere decir que son preparados para ser explotados o usados. No obstante, la producción y extracción de los datos cuenta con parámetros

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

definidos para que estos sean de calidad, es necesario realizar una revisión y verificación. Para el desarrollo de este componente se demanda ejecutar las siguientes actividades:

- A.** *Revisar las bases de datos:* consiste en realizar una verificación a la base de datos, una vez ha sido culminada, con el fin de garantizar la implementación del componente de extracción y producción y que los datos correspondan a lo establecido en la matriz de metadatos y en la de datos maestros; como resultado de esta revisión se obtendrá un hallazgo de errores cometidos.
- B.** *Definir las acciones de mejora:* en caso de encontrar errores, se debe definir el procedimiento para solventar las novedades halladas; esto, con el ánimo de definir qué hacer con los campos que quedaron vacíos, duplicados o erróneos dentro de las bases de datos. Se espera, entonces, transformar los datos para alcanzar el formato requerido, normalizar los valores, eliminar (si es necesario) los datos duplicados, promediar los datos o validarlos.
- C.** *Ajustar las bases de datos:* conforme con la revisión y hallazgos encontrados en las bases de datos, y de acuerdo con las acciones de mejora establecidas, se deberán adaptar las bases de datos, de modo que los mismos sean de calidad y se encuentren preparados para generar valor en la estrategia de explotación. Cabe resaltar que esta actividad de ajuste de la base de datos puede ser ejecutada mediante herramientas configuradas para ello.

7.3.2 Valoración de datos

La valoración es la estrategia mediante la cual un grupo interdisciplinario de profesionales analiza la producción de los datos, su necesidad e importancia para las entidades y la comunidad en general; es transversal, lo que significa que está pensada para apoyar las demás estrategias, cuyo propósito es garantizar una adecuada gestión archivística de los datos. Para tal logro, se desarrollaron los componentes sobre analítica y borrado seguro de los datos.

7.3.2.2 Componente: analítica de datos

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

Este componente corresponde al estudio de los datos, enfocado en su análisis con fines de valoración. Para su logro, se determina la conformación de un grupo interdisciplinario, los métodos y técnicas para el estudio y, por último, la formulación de tablas de retención de datos, entendidas como ese instrumento de la gestión archivística en el cual se define la disposición final de los datos y sus tiempos de retención. Para el desarrollo de este componente es necesario ejecutar las siguientes actividades:

- A.** *Conformar un grupo interdisciplinario:* consiste en definir un grupo encargado de la valoración de los datos, conformado por profesionales en archivística y tecnologías de la información, así como personal de diferentes áreas estratégicas de las entidades relacionadas con su misión, con diversas habilidades, perspectivas y experiencias, el contexto en que se produjeron los datos y las funciones que cada área ejecuta diariamente. Con lo anterior se busca que este componente permita la ejecución de la estrategia de valoración de manera transversal e interdisciplinaria.
- B.** *Definir los métodos y técnicas para el análisis:* con base en las necesidades identificadas de valoración de los datos, se pasa a especificar un método para su análisis, de acuerdo con el tipo de dato —ya sea cualitativo, estadístico u otro— así como las técnicas para la ejecución de este método —como el análisis descriptivo, estadístico, entre otros— que posibiliten analizar los datos con miras a definir los valores que estos tienen para las entidades.
- C.** *Analizar el contexto institucional de la entidad:* consiste en definir la trayectoria, funciones, misión, visión, así como el entorno de las entidades, como una de las primeras actividades a ejecutar, para determinar los valores de los datos, acorde con su relevancia para las entidades y la comunidad en general.
- D.** *Definir los valores de los datos:* una vez se define el contexto de las entidades, así como el desarrollo de sus funciones, el grupo interdisciplinario podrá definir qué valores serán asignados a los datos. Valga diferenciar entre aquellos primarios: asociados y vinculados con la misión de la entidad, es decir, que otorgan valor a los datos, en relación con las funciones de la entidad y, los secundarios, asociados con el reconocimiento de la memoria institucional,

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

esto es, datos que por sus características se pueden vincular al desarrollo histórico y al aporte social de la entidad.

- E. *Elaborar la tabla de retención de datos*: este instrumento debe estar en concordancia con la matriz de metadatos y la de datos maestros, ya que con él se identifican los datos generados por las entidades, en razón de sus funciones, y establecen los tiempos de retención y disposición final, y se define si la base de datos es eliminada o se conserva de forma permanente, en línea con los valores definidos para los datos. Esta tabla responde a las agrupaciones establecidas en la estrategia de producción y extracción de datos, en sintonía con las funciones que les dieron origen. Así, se aseguran unas categorías ya identificadas.

Es importante la vinculación de las tablas de retención de datos y las tablas de retención de documentos, ya que en ambos casos se está hablando del producto de la ejecución de las funciones de las entidades; el resultado será obtener datos y documentos, resultantes de una misma actividad o función, en consecuencia, con valores similares o iguales; por ejemplo: el proyecto o contrato para censar la población, de lo cual se derivan una cantidad de documentos que muestran la ejecución del mismo —actas de inicio, presupuesto y otros—; también la base de datos, producto de este proyecto o contrato, y los informes para los que se tuvieron en cuenta bases de datos como insumo de los mismos. Con todo, es necesario mantener un vínculo entre dichos instrumentos, ya que esto permite una mejor comprensión de los datos y los documentos relacionados, así como contar con datos fiables y utilizables. De esta manera los usuarios podrán contar con información complementaria y así potencializar el uso de los datos y documentos, como complemento y en conjunto.

7.3.2.4 *Componente: borrado seguro de datos*

Como su nombre lo indica, es el proceso de borrado de las bases de datos, previamente identificadas y seleccionadas, y establecido en las tablas de retención de datos. En su proceso, es necesario ejecutar las siguientes actividades:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- A. *Socialización del proceso de eliminación o borrado:* consiste en dar a conocer aquellos datos o bases de datos que serán eliminados al interior de las entidades, con el fin de que la eliminación de los datos sea transparente. Para su consecución es necesario elaborar un acta para socializar este lineamiento y donde se comunique a los interesados la disposición final de estos datos, de acuerdo con las tablas de retención de datos. Finalmente, las objeciones, inquietudes o solicitudes, frente al proceso de borrado de los datos, serán recibidas en un término de días que establezca la entidad para ello.

- B. *Definir la técnica para la eliminación:* establecer cómo se eliminarán o borrarán los datos, así como los responsables de esta actividad, que pueden ser los representantes del grupo interdisciplinario (profesional en sistemas o personal del área de información y científico de datos). La técnica elegida debe estar concebida para la eliminación total de los datos ya identificados; además, que estos no queden almacenados en algún respaldo, ya que la finalidad de eliminar los datos es la optimización de recursos y del espacio de almacenamiento de aquellos que ya perdieron valor para la entidad y para la comunidad en general.

- C. *Eliminar o borrar los datos:* una vez ejecutadas las actividades anteriores, se podrá proceder con la eliminación de los datos, de acuerdo con las modificaciones que se hayan dado luego de la socialización de esta decisión; además, se deberá informar al interior de las entidades y la comunidad en general, una vez se ejecute esta actividad. Finalmente, es necesario verificar la ejecución de esta actividad con el fin de evitar el almacenamiento de datos en respaldo.

7.3.3 *Preservación de datos*

Garantizar la preservación de los datos y su acceso a lo largo del tiempo, es el objetivo de este componente. En efecto, se pretende contar con datos de calidad, fiables, auténticos y completos, razón por la cual se llevarán a término los componentes de almacenamiento y protección de los datos.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

7.3.3.1 Componente almacenamiento

El almacenamiento con miras a la preservación de los datos se hará en repositorios establecidos para tal fin, ya que estos permiten mantenerlos disponibles a lo largo de su ciclo de vida. Para el desarrollo de este componente se plantea la definición del repositorio, la normalización de los datos y la actualización de servidores o equipo, así como los formatos. Para el fin de este componente es necesario ejecutar las siguientes actividades:

- A. *Definir el repositorio de datos:* seleccionar el espacio centralizado de acuerdo con las necesidades de almacenamiento, teniendo en cuenta el volumen de los datos en las entidades. Una vez se cuente con el repositorio, es necesario realizar vigilancia tecnológica sobre esta herramienta, de tal forma que se haga un monitoreo permanente de accesibilidad de los datos, con el fin de mitigar su pérdida. Para ello, es necesario estar en constante actualización frente a las tecnologías usadas y plantear acciones de mejora oportunamente.
- B. *Normalizar los datos:* esta actividad consta de identificar aquellos datos maestros y que estos respondan a los lineamientos sobre estructura definidos en la matriz de metadatos. En esta línea de sentido, se espera que estos sean abiertos y usables. Para alcanzar la normalización, es necesario definir todos los sistemas de gestión y almacenamiento de datos con que cuentan las entidades, de los cuales se deben describir y garantizar su relacionamiento con los demás sistemas; es esencial, pues, definir los formatos compatibles de las bases de datos para los sistemas identificados, procurando que sean accesibles en cada uno de ellos.
- C. *Actualizar servidores o equipos y formatos:* esta actividad responde a la necesidad de renovación de la tecnología, teniendo en cuenta las dinámicas actuales y avances en esta materia, para lo cual se requiere definir y contar con acciones como migración de los datos: cambiarlos de un sistema a otro; *refreshing*: mover los datos de un medio digital a otro; intercambio de datos: trasladar los datos de un servidor a otro y demás alternativas que permitan mantenerlos accesibles y usables.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

7.3.3.2 Componente protección de datos

Este componente se enfoca en garantizar la seguridad de los datos, ajustado a los diferentes tipos que se tengan en las entidades, identificados partiendo de su nivel de acceso, como los personales que cuentan con niveles de reserva, datos públicos o privados. Su proceso demanda la ejecución de las siguientes actividades:

- A. *Realizar copias de seguridad*: consiste en hacer un respaldo de los datos en el repositorio o unidad alterna que permita recuperar los datos ante eliminaciones no controladas, manipulación inadecuada y ataques cibernéticos; así mismo, se deberá establecer un cronograma para realizar este respaldo de los datos. Para el desarrollo de esta actividad, se debe definir un responsable, conocedor del repositorio, la periodicidad con que se deban realizar e, incluso, su incremento, siempre teniendo en cuenta las necesidades de continuidad del negocio.
- B. *Elaborar las tablas de control de acceso*: definir un instrumento donde se tomen aquellos datos que por sus características tengan alguna restricción para su acceso, y se definan aquellos usuarios que puedan o no acceder a los datos de tipo público, privado, personales, impersonales, clasificados, reservados y confidenciales. Para ello se recomienda tener en cuenta las funciones del personal y su tipo de contrato y, para los usuarios externos, es menester caracterizarlos y definir quiénes y bajo qué condiciones pueden acceder a los datos.
- C. *Garantizar la seguridad de los datos*: el desarrollo de esta actividad se realizará, principalmente, con la implementación de las tablas de control de acceso, ya que esto consentirá mantener los datos bajo el uso y acceso de personas autorizadas; además, para la seguridad de los datos también deben tenerse en cuenta las actividades del componente de almacenamiento y los riesgos a los que se ven expuestos los datos, con el fin de mitigarlos y mantenerlos protegidos.

Las estrategias para la gestión de datos desde la perspectiva archivística, en todo caso, se proponen como una solución a la problemática planteada en esta investigación, ya que estas evitarán la conservación de datos sin ninguna utilidad y, en este sentido, posibilitarán la optimización de

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

recursos, el acceso ágil a los datos y su disponibilidad. Conviene precisar que para este caso se propusieron y desarrollaron tres estrategias: generación, valoración y preservación de datos, poniendo en perspectiva las oportunidades que se presentaron para la gestión documental, luego de desarrollar el comparativo con la gestión de datos.

8. Conclusiones

Para comprender la gestión de datos desde la perspectiva archivística, fue necesario abarcar y analizar la gestión de datos y la gestión documental. En este primer momento, se lograron identificar aspectos en común de acuerdo con la bibliografía analizada, es decir, los conceptos, los tipos de datos y documentos que se contemplan en cada uno de estos, así como sus principios y procesos; esto, a su vez, permitió tener las bases y el conocimiento suficientes para el desarrollo de los demás objetivos planteados en esta investigación.

Con las características de la gestión de datos y la gestión documental, se encontraron una serie de similitudes, algunas bastante evidentes, como el uso de metadatos durante la gestión de datos y de los documentos electrónicos, así como el proceso de gestión de documentos y contenidos, situación que permitió vislumbrar las relaciones a que habría lugar. Durante el desarrollo del segundo objetivo, al comenzar con la comparación y la identificación de similitudes y diferencias, no solo se vio la necesidad, sino que también se logró definir aquellos puntos de complementariedad, es decir, aquellos elementos de la gestión documental que aportan a una gestión de datos desde una perspectiva archivística. Ahora bien, se podrán realizar investigaciones comparativas basadas en los diferentes modelos existentes para la gestión de datos y la gestión de documentos, que brinden una perspectiva complementaria a la investigación realizada sobre similitudes, diferencias y oportunidades que puedan surgir.

En este sentido, se proponen como estrategias para la gestión archivística de datos en Colombia, la generación de datos, en virtud a que en esta última se lograron definir elementos sobre la planeación, la producción, la extracción y la limpieza de estos; la segunda estrategia sobre la valoración de los datos, permite que estos sean analizados para definir sus tiempos de retención, su disposición final y borrado seguro, logrando que sean conservados de acuerdo con las necesidades, funciones y misión de cada entidad, en el entendido de que cada caso es diferente. La última de las estrategias es la preservación de los datos, teniendo presente que los datos se gestionan y almacenan

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

en herramientas tecnológicas que cambian constantemente. En concreto, la estrategia en mención busca solventar y minimizar el riesgo de pérdida de información por deterioro, obsolescencia y manipulación inadecuada de las herramientas.

Expuesto así, la producción masiva de datos puede ser controlada mediante la implementación de las estrategias para la gestión archivística de los datos, planteadas en el desarrollo de esta investigación, ya que con estas se busca administrar los datos durante todo su ciclo de vida; asimismo, se garantiza la gestión y uso de los datos en contribución al fin último de los archivos, el acceso a la información; lo planteado, con el propósito de que los usuarios obtengan respuesta a sus requerimientos de una forma eficaz y exacta. Es así como se da respuesta a la pregunta planteada y al objetivo general, acerca de cómo gestionar los datos desde la perspectiva archivística en Colombia.

En cuanto a las limitaciones durante el proceso de investigación, fue la poca bibliografía disponible lo que dificultó la discusión en el ámbito académico, y la definición de su viabilidad desde el inicio; sin embargo, este obstáculo fue asumido como oportunidad, al realizar y desarrollar una investigación bajo una temática específica poco explorada. Se evidenció, además, la necesidad de que el archivista, que cuenta con los estudios y experiencia en el manejo de documentos, tenga el apoyo de diferentes profesiones —como ingenieros y científicos de datos—, pues su competencia es de alta valía al contar con el conocimiento en las tecnologías de la información y de manejo de datos. A propósito, es claro que se trata de una debilidad, ya que la investigación fue desarrollada por una archivista; no obstante, se puede convertir en una oportunidad, por cuanto abre la posibilidad para conformar grupos de carácter interdisciplinario y aportar a estos temas de investigación.

Finalmente, es válido resaltar que la gestión archivística de datos es un desarrollo necesario en Colombia, puesto que se constituye en un aporte complementario para la gestión documental y, por ende, para la disciplina archivística, de manera que se tengan nuevos retos y que surjan una serie de contribuciones entre ambas partes. Está claro que esta investigación planteó una serie de estrategias; no obstante, es necesario realizar cambios y ajustes pertinentes (adaptar los instrumentos y procesos de la gestión documental a las necesidades de los datos), desde lo académico y lo investigativo, y así expandirlo a lo aplicable, lo laboral, para obtener una formulación completa e íntegra. Todo ello demanda una implementación de las estrategias planteadas, más aún, una complementariedad y mejora de estas, no solo con más perspectivas, sino

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

también con otras herramientas que ayuden y faciliten esta gestión, como la explotación de datos como un posible servicio en unidades de información que den repuesta exacta a los requerimientos de los usuarios.

Recogiendo todo lo planteado, es menester afirmar que la actualización y avance del entorno, en general, obliga a una constante mejora de estrategias, procesos y herramientas a utilizar, con el fin de garantizar la gestión archivística de los datos y, en coherencia, garantizar un acceso ágil y oportuno a información de calidad.

9. Referencias bibliográficas

- Aguilera Murguía, R. (2012). *Archivística integrada*. México, Escuela Mexicana de Archivos A. C. (EMAAC).
- Alonso Varela, Lucía; Noble Moreira, Luis Daniel e Saraiwa Cruz, Ignacio. 2015. El concepto de paradigma en la Archivística como problema epistemológico. *Palabra Clave* (La Plata), vol. 4, n° 2, p. 109- 128. Disponible en:
<http://www.palabraclave.fahce.unlp.edu.ar/article/view/PCv4n2a04>.
- Amutio G., M. Á. (2021). La orientación al dato, nuevos desafíos. *Tábula*, (23), 99-126.
<https://doi.org/10.51598/tab.844>
- Archivo General de la Nación, (2006) Acuerdo 027 de 2006 Por el cual se modifica el Acuerdo No. 07 del 29 de junio de 1994.
- Archivo General de la Nación, (2012). *Guías de Cero Papel de la Administración Pública; Definiciones, características y metadatos de los expedientes en soportes electrónicos*, de <http://preservandoparaelfuturo.org/>
- __. (2014). *Manual para la Implementación de un Programa de Gestión Documental*.
[https://www.archivogeneral.gov.co/caja_de_herramientas/docs/2.%20planeacion/MANUALES/PArchivo General de la Nación. \(2015\). Gestión documental y gobierno electrónico: problemas, retos y oportunidades para los profesionales de información.](https://www.archivogeneral.gov.co/caja_de_herramientas/docs/2.%20planeacion/MANUALES/PArchivo%20General%20de%20la%20Nacion%20(2015).Gestion%20documental%20y%20gobierno%20electronico%20problemas,%20retos%20y%20oportunidades%20para%20los%20profesionales%20de%20informacion.pdf)
- __. (5 de julio de 2018). *Requisitos mínimos de digitalización*.
- __. (9 de mayo de 2019). *Modelo de requisitos para la implementación de un sistema de gestión de documentos electrónicos de archivo*.
- __. (2019). *Modelo de requisitos para la implementación de un SGDEA*.
- __. (2020). *Guía de Implementación de un Sistema de Gestión de Documentos Electrónicos de Archivo*.
https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/V16_Guia_SGDEA.pdf
- __. (2020). *Guía de implementación de un Sistema de Gestión de Documentos Electrónicos de Archivo SGDEA*.
- __. (2020). *Modelo de Gestión Documental y Administración de Archivos*.
https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/V3_MGDA-min.pdf

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- . (2020). Guía de anonimización de datos estructurados: Conceptos generales y propuesta metodológica.
- Arriagada-Benítez, M. (2020). Ciencia de Datos: hacia la automatización de las decisiones. *Ingeniare. Revista chilena de ingeniería*, 28(4), 556-557.
- Betancur-Roldán, M. C. y Jaramillo, O. (2021). Aproximación a la evolución del concepto de archivística. *Revista Interamericana de Bibliotecología*, 44(1)
<https://doi.org/10.17533/udea.rib.v44n1eIn1>
- Ruesta, C. B. (2018). Transformación digital desde la perspectiva de la gestión documental. *RUIDERAE: Revista de Unidades de Información*, (13).
- Bustelo Ruesta, C. (2021). Gobernanza de datos y gestión de documentos. *Tábula*, (24), pp. 65-79
- Capurro, R. (2014). Pasado, presente y futuro de la noción de información. *Logeion: filosofia da informação*, 1(1), 110-136.
- Caraballo Pérez, Y., & Ramírez Céspedes, Z. (2007). El ciclo de vida de la información en la Web. *Ciencias de la Información*, 38(1-2), 31-40. [fecha de Consulta 22 de Agosto de 2023]. ISSN: 0864-4659. Recuperado de:
<https://www.redalyc.org/articulo.oa?id=181414865004>
- Congreso de la República de Colombia (2000). Ley 594 de 2000: Por medio de la cual se dicta la Ley General de Archivos y se dictan otras disposiciones.
<https://normativa.archivogeneral.gov.co/ley-594-de-2000/>
- . (2008). Ley 1266 "Por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones". Recuperada de:
<https://intranet.secretariajuridica.gov.co/transparencia/marco-legal/normatividad/ley-1266-2008#:~:text=Descripci%C3%B3n%3A,y%20se%20dictan%20otras%20disposiciones.>
- . (2012). Ley estatutaria 1581. Por la cual se dictan disposiciones generales para la protección de datos personales. Recuperada de:
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>
- . (2014). Ley 1712. Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=56882>
- Mundet, C., & Archivística, J. M. (2011). Principios, términos y conceptos fundamentales. MUNDET, JRC Administración de documentos y archivos: textos fundamentales. *Guillomía Comunicación Gráfica*, 15-35.
- Cunningham, A. (2021). ¿Cómo de lleno está el vaso? Cambios y desafíos para los profesionales de los documentos frente a la transformación digital en la era de los datos. *Tábula*, (24), pp. 25-40.
- Departamento de Políticas Digitales y Administración Pública (2020). Decreto 76/2020, de 4 de agosto, de Administración digital.
https://noticias.juridicas.com/base_datos/CCAA/672737-d-76-2020-de-4-ago-ca-catalunaadministracion-digital.html
- Departamento Nacional de Planeación, Ministerio de Tecnologías de la Información y las Comunicaciones, Superintendencia de Industria y Comercio. (17 de abril de 2018). Documento CONPES 3920: política nacional de explotación de datos (*big data*).

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- Domingo, G., Buccella, A. y Cechich, A. (2007). Un marco de trabajo para analizar y mejorar la calidad de datos dentro de su ciclo de vida. En *XIII Congreso Argentino de Ciencias de la Computación*.
- García-Alsina, M. (2017). Transversalidad y gestión: documentos y datos al servicio de la toma de decisiones y la transparencia. *RUIDERAe: Revista de Unidades de Información*, (11).
- García-Morales, E. (2013). Gestión del ciclo de vida de datos y documentos: acercando posiciones. *Anuario ThinkEPI*, 7, 98-100.
- Global Data Management Community. (2017). DAMA-DMBOK: Data Management Body of Knowledge (2ª. ed.). Technics Publications. Technics Publications.
- Grant, R. (2017). Recordkeeping and research data management: a review of perspectives. *Records Management Journal*.
- Guzmán, W. C. V., Cadena, A. G. M., Escalante, A. M. O. y Hivon, M. S. (2021). Importancia del *big data* en un gestor documental para las entidades públicas de Colombia. *SIGNOS Investigación en Sistemas de Gestión*, 13(1), 1-21.
- Hemphill, L., Hedstrom, M. L. & Leonard, S. H. (2021). Saving social media data: understanding data management practices among social media researchers and their implications for archives. *Journal of the Association for Information Science and Technology*, 72(1), 97-109.
- Hernández Orallo, J. (2004). *Introducción a la Minería de Datos*. Editora E-papers.
- Herrera, A. H. (2018a). El servicio y el uso: del documento a la información. *Revista TRIA*, (22), 273-288.
- ___ (2018b). El servicio y el uso: del documento a la información. *Revista TRIA*, (22), 273-288.
- ___ (2019). Terminología archivística: Conceptos y uso. *Revista TRIA*, (23), 63-76.
- i Serra, L. E. C. (2016). A la preservación de datos... ¡ y más allá!. *Boletín del Archivo General de la Nación*, 8(09), 157-187.
- Ibáñez Pascual, A. (2021). Gestión y visualización de datos abiertos gubernamentales. El caso de los datos sobre la pandemia del COVID-19 en Castilla y León. *Tábula*, (24), pp. 183-198
- Instituto Colombiano de Normas Técnicas. (2013). ISO - NTC 30300. Información y documentación. Sistemas de gestión para registros. Fundamentos y vocabulario. <https://ecollection-icontec-org.udea.lookproxy.com/normavw.aspx?ID=6230>
- ___ (2017). ISO - NTC 15489. Información y documentación. Gestión de registros. Parte 1: Conceptos y principios. <https://ecollection-icontec-org.udea.lookproxy.com/normavw.aspx?ID=6163>
- Iramain, J. L. (2000). Una historia del concepto información de la causa formal al dato (y vuelta). *Communication & Society*, 13(1), 91-114.
- Jefatura del Estado. España. (6 de diciembre de 2018). Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. «BOE», (294), de Referencia: BOE-A-2018-16673
- Lecanda, R. Q. y Garrido, C. C. (2002). Introducción a la metodología de investigación cualitativa. *Revista de Psicodidáctica*, (14), pp. 5-40.
- López C., A. M., Requejo Zalama, J. (2021). Estrategias y modelos de gestión de datos archivísticos. *Tábula*, (24), pp. 21-4.
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M. & Conrad, M. (2018). *Archival records and training in the age of big data*. In *Re-Envisioning the MLS: Perspectives on the future of library and information science education*. Emerald Publishing Limited.

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- Martínez García, L. (2021). ¿De verdad estamos capacitados para archivar la “república de los datos” en un entorno digital? *Tábula*, (24), pp. 81-95.
- Martínez M., C. (2021). Ciencia de Datos versus Ciencia de la Información: informetría y análisis textual. *Scire: Representación y organización del Conocimiento*, 27(2), 31-42.
<https://doi.org/10.54886/scire.v27i2.4731>
- Ministerio de Tecnologías de la Información y las Comunicaciones (2019). Guía técnica de Información: Administración del dato maestro.
https://www.mintic.gov.co/arquiturati/630/articles-9254_recurso_pdf.pdf
- ___ (2019). Guía Técnica de Información: Gobierno del dato.
https://www.mintic.gov.co/arquiturati/630/articles-9258_recurso_pdf.pdf
- Moro-Cabero, M. (2021). Datos y archivos: retos para una profesión agitada por la transformación digital. *Acervo*, 34(3), pp. 1-23.
- Mundet, C. R. U. Z., & Archivística, J. M. (2011). Principios, términos y conceptos fundamentales. MUNDET, JRC Administración de documentos y archivos: textos fundamentales. Madrid: Guilmoría Comunicación Gráfica, 15-35.
- Musiño, C. M. (2021). Ciencia de datos versus ciencia de la información: informetría y análisis textual. *Scire: representación y organización del conocimiento*, 27(2), 31-42.
- Online etymology dictionary. (2022). *Definición de dato*.
<https://www.etymonline.com/search?q=dato>
- Pérez Borges, A., (2009). Relación fuente - recurso de información - documento. *Biblios*, (37).
<https://www.redalyc.org/articulo.oa?id=16119333004>
- Pérez, Y. C. & Céspedes, Z. R. (2007). El ciclo de vida de la información en la Web. *Ciencias de la Información*, 38(1-2), 31-40.
- Ponjuan Dante, G. (Diciembre de 2005). Gestión documental, gestión de información y gestión del conocimiento: evolución y sinergias. Comunicación preliminar. *Ciencias de la Información*, 36(3).
- Presidencia de la República de Colombia (2012). Directiva Presidencial 04. Eficiencia administrativa y lineamientos de la política cero papel en la administración pública.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=50155>
- ___ (2015). Decreto 1080. Por medio del cual se expide el Decreto Único Reglamentario del Sector Cultura.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=76833>
- Quecedo, R., y Castaño, C. (2002). Introducción a la metodología de investigación cualitativa. *Revista de Psicodidáctica*, (14), 5-39. <https://www.redalyc.org/articulo.oa?id=17501402>
- Rincón, R. D. (1998). Los indicadores de gestión organizacional: una guía para su definición. *Revista Universidad EAFIT*, 34(111), 43-59.
- Ruiz, F. F. (1999). Archivística, archivo, documento de archivo... Necesidad de clarificar los conceptos. In *Anales de documentación* (Vol. 2, pp. 103-120). Facultad de Comunicación y Documentación y Servicio de Publicaciones de la Universidad de Murcia.
- Salvador Oliván, J. A. (2002). Formación en recuperación de información: análisis de los cursos y asignaturas en las escuelas de biblioteconomía y documentación de Norteamérica y España. *Documentación de las Ciencias de la Información*, 25, 189-215.
- Serra S., J. (2021a). Conferencia Inaugural del Segundo Semestre 2021. La gestión documental orientada a los datos. Transmitida por: Ciudad de la Investigación de la Universidad de Costa Rica, San José, Costa Rica. En:

La gestión de datos en Colombia: un estudio desde la perspectiva archivística

- <https://www.facebook.com/ArchivisticaCR/videos/conferencia-inaugural-ii-semester-la-gesti%C3%B3n-documental-orientada-a-datos/251510053492031/>
- ___ (2021b). Gestión documental datificada. En el consultor de los ayuntamientos y juzgados (ed.). *La gestión del documento electrónico* (pp. 511-526). Editorial
- Serra Serra, J. (2021c). Los datos: ¿una amenaza? Gestión de documentos en la era de los datos. *Tábula*, (24), pp. 43-63.
- Subdirección General de los Archivos Estatales, España. (1995). *Diccionario de Terminología Archivística*.
- Universidad de Antioquia. (2022). Escuela Interamericana de Bibliotecología: Maestría en Ciencia de la Información. <https://acortar.link/5opj5w>
- Vargas Guzmán, W. C., Moreno Cadena, A. G., Oñate Escalante, A. M. y Sanabria Hivon, M. (2020). Importancia del *big data* en un gestor documental para las entidades públicas de Colombia. *SIGNOS, Investigación en Sistemas de Gestión*, 13(1), pp. 1 - 21.
- Wu, X., Zhu, X., Wu, G. Q. & Ding, W. (2013). Data mining with *big data*. *IEEE transactions on knowledge and data engineering*, 26(1), 97-107.
- Yeo, G. (2018). Records and Data. In *Records, Information and Data: Exploring the Role of Record Keeping in an Information Culture* (pp. 105-128). Facet.
doi:10.29085/9781783302284.006
- Zea Elizalde, E. A. y Pulido Daza, N. J. (2016). Estado del arte de un sistema de gestión documental electrónica desde la norma ISO 15489-1: 2006. *Códices*, 12(2), 43-77.