GUÍA PARA LA ANONIMIZACIÓN DE DATOS E INFORMACIÓN NO ESTRUCTURADA:

ESTÁNDARES Y LINEAMIENTOS TÉCNICOS

La equidad es de todos
Prosperidad Social
Centro Nacional de Memoria Histórica

GUÍA PARA LA ANONIMIZACIÓN DE DATOS E INFORMACIÓN NO ESTRUCTURADA: ESTÁNDARES Y LINEAMIENTOS TÉCNICOS

Establecimiento público adscrito al Departamento para la Prosperidad Social (DPS)

Rubén Darío Acevedo Carmona
Director general

Marcela Rodríguez Vera
Directora Archivo de los Derechos Humanos

Autores
Centro Nacional de Memoria Histórica: John Garzón, Valeria Eraso Cruz, Natasha Eslava Vélez, Marcela Rodríguez Vera

Archivo General de la Nación: Laura Sánchez Alvarado, Diana Paola Vásquez Zea, Diana Monroy García

Revisó
Lilian Lizbeth Barrientos
Presidente Comité Nacional de Memoria del mundo - Guatemala

John Francisco Cuervo Alonso
Presidente de la Sociedad Colombiana de Archivistas

Documento aprobado por el Comité Editorial en mayo 24 de 2022

Revisión de textos
Natasha Eslava Vélez, Alejandro Triana Laverde

Diagramación y gráficas
Gabriel Gómez Vargas

Fotografía
Freepik.es

ISBN
978-628-7561-24-3

Bogotá D. C., Colombia – 2022
Primera versión 2022

CentroMemoriaH @CentroMemoriaH

http://archivodelosddhh.gov.co/
https://centrodememoriahistorica.gov.co/

TABLA DE CONTENIDO

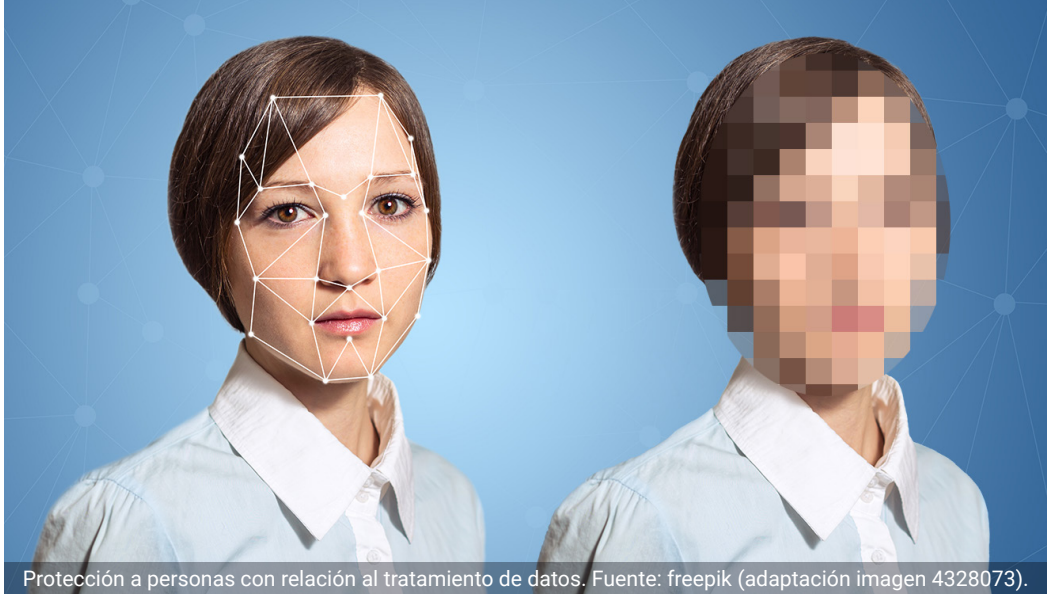
1. INTRODUCCIÓN

En el marco del Programa de Derechos Humanos y Memoria Histórica de que trata el artículo 144 de la Ley 1448 de 2011, la Dirección de Archivo de los Derechos Humanos del CNMH trabaja en función de reunir, ordenar, clasificar y describir los documentos relativos a la violación de los derechos humanos y el DIH, con el propósito de garantizar su custodia, preservación y uso por parte de investigadores y público en general, que a su vez sirven como plataforma de apoyo, gestión, intercambio y difusión en los temas de memoria histórica, promoviendo la participación de las víctimas, con enfoque diferencial y, además como espacio de apoyo a las entidades públicas y privadas en el marco de las iniciativas ciudadanas en temas de memoria histórica.

Para ello, ha integrado un archivo con documentos de las violaciones ocurridas con ocasión del conflicto armado interno, así como documentación sobre procesos similares en otros países, ha recopilado testimonios orales, escritos y de toda índole, por medio de los ejercicios investigativos realizados por la entidad, o que le han sido allegados por las organizaciones sociales de derechos humanos, con el fin de reunirlos, preservarlos y garantizar su custodia. No obstante, estos documentos se enfrentan a desafíos de interpretación, tanto a la hora de su solicitud, como sobre qué tipo de información debe ser pública y cómo resolver casos en los que se involucre información reservada o datos personales o sensibles y por ello es pertinente conocer algunas prácticas sobre cómo operar en la eventualidad en que se presenten estas situaciones, con miras a lograr una mayor protección y acceso a la información recopilada.

En vista a lo anterior, el presente documento se armoniza con la Política de seguridad y privacidad de la información del CNMH, mediante lineamientos enfocados a la seguridad de la información que deben conocer, acatar y cumplir todos los funcionarios, contratistas, personal en comisión, visitantes y terceros que presten sus servicios o tengan algún tipo de relación con el CNMH.

En Colombia, la Ley general de protección de datos personales (Congreso de la República, 2012) establece los principios aplicables a las actividades de tratamiento de datos personales para garantizar el derecho fundamental de hábeas data de las personas. Así, dentro de ese listado de principios, previsto en el artículo 4 de la Ley 1581 de 2012, se incluye el de circulación restringida y seguridad, que señala que los datos personales, salvo la información pública, no pueden estar disponibles en internet u otros medios de divulgación o comunicación masiva, y lo podrán estar cuando el acceso sea técnicamente controlable para brindar un conocimiento restringido de la información. Esta información debe



Protección a personas con relación al tratamiento de datos. Fuente: freepik (adaptación imagen 4328073).

conservarse con las medidas técnicas, humanas y administrativas necesarias para otorgarle seguridad evitando su adulteración, pérdida, consulta, uso o acceso no autorizado o fraudulento.

Para ello, es necesario realizar una técnica denominada anonimización, que se aplica a los datos personales para obtener una desidentificación no reversible, es decir, lograr que, no se identifique posteriormente a las personas cuya información es recolectada y usada.

Según la Guía de instrumentos de información pública (2016), es importante tener en cuenta los siguientes aspectos, del proceso de anonimización:

> El proceso de anonimizar información no sólo implica la eliminación de las variables de identificación directa de la unidad de observación (por ejemplo, la cédula o el NIT de una empresa), sino que se deben realizar procedimientos adicionales para garantizar la confidencialidad de los datos. (Secretaría de Transparencia de la Presidencia de la República, 2016)
>
> Se debe valorar el "riesgo de revelación individual", es decir, la probabilidad que tiene una observación de ser descubierta a partir de características que contiene la información. (…) Esta valoración se puede realizar de manera rigurosa mediante la programación de algoritmos en programas estadísticos que arrojan la probabilidad exacta del riesgo de revelación individual, o analizando una por una las variables e identificando aquellas que aumentan este riesgo. (P.65)

En ese orden y con el propósito de dar cumplimiento a la normatividad en materia de protección de datos personales y, a la vez, lograr un aprovechamiento de datos en el Estado y que se garantice el derecho de acceso a la información pública, además de proporcionar una orientación metodológica y una orientación sobre las técnicas para realizar procesos de anonimización de datos e información producidos o gestionados por entidades públicas y privadas, o que hagan parte de la justicia transicional, se presenta la *guía para la anonimización de datos e información*, que tiene como fundamento las disposiciones jurídicas nacionales e internacionales en materia de privacidad, protección de datos personales, acceso y transparencia a la información pública y su aplicación le corresponde a las entidades públicas y a las empresas privadas que desarrollan funciones públicas y que intercambian información (COMPES 3920, 2018).

De igual manera, su objeto es la protección de cualquier información producida, gestionada o recolectada por esas entidades, públicas o privadas, que contenga datos personales o información bajo las siguientes premisas:

I. Protección de derechos: garantía en el tratamiento de datos personales y privados.

II. Transparencia y datos abiertos: el derecho de acceso a la información pública, la transparencia activa y pasiva.

III. Acceso e interoperabilidad: condiciones mínimas para el uso de datos.

IV. Eficiencia administrativa.

V. Reportes de información.

En la primera parte de la guía de estándares y lineamientos técnicos para la anonimización de datos e información personal se encuentran algunas definiciones extraídas directamente de la normatividad nacional e internacional de términos que tienen que ver directamente con el objeto de la Guía y se señalan los principios aplicables en los procesos de anonimización.

En la segunda parte, se presentan las técnicas de anonimización describiendo sus características esenciales y entregando ejemplos explicativos que facilitan la comprensión de estas.

En la parte final se hace referencia a los riesgos asociados al proceso de anonimización de información personal y se incluye una tabla de resumen de técnicas de anonimización y un normograma que facilita la consulta.

Es importante advertir que los estándares, lineamientos y las técnicas para llevar a cabo los procesos de anonimización deben ser actualizados y evaluados regularmente para prever y corregir los factores de riesgo.

Anonimización de la identidad personal. Fuente: freepik (adaptación imagen 5604325).

2. GLOSARIO

Anonimización: el proceso por el cual la información de identificación personal se modifica de forma irreversible de tal manera que no se pueda identificar, directa o indirectamente, ya sea por sus propios medios o en colaboración con algún tercero, a la persona asociada a dicha información de identificación personal. (Estandar ISO / IEC 29100:2011, 2011).

Archivo: es el conjunto de documentos, sea cual fuere su fecha, forma y soporte material, acumulados en un proceso natural por una persona o entidad pública o privada, en el transcurso de su gestión, conservados respetando aquel orden para servir como testimonio e información a la persona o institución que los produce y a los ciudadanos, como fuentes de la historia. También se puede entender como la institución que está al servicio de la gestión administrativa, la información, la investigación y la cultura. (Ley 1712, 2014. Artículo 6).

Base de datos: conjunto de datos organizado de tal modo que permita obtener con rapidez diversos tipos de información. (RAE, 2017). Conjunto organizado de datos personales que sea objeto de tratamiento. (Ley 1581, 2012. Artículo 3).

Conjunto de datos: unidad mínima de información sujeta a carga, publicación, transformación y descarga. (Guía de datos abiertos en Colombia, 2016, p. 17).

Dato: representación primaria de variables cualitativas y cuantitativas que son almacenables, transferibles, pueden ser visualizadas, controladas y entendidas [...] Los datos que se perciben por los sentidos humanos, de manera continua y sin interrupciones, se denominan analógicos. Cuando son interpretados mediante codificación binaria, se trata de datos digitales. Estos últimos 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. (Consejo Nacional de Política Económica y Social - CONPES, 2018, p. 25).

Dato estructurado: están organizados conforme a un modelo o esquema. Se almacenan en forma tabular y algunas veces su estructura también incluye la definición de las relaciones entre ellos. Típicamente están representados en bases de datos que hacen parte del funcionamiento de sistemas de información. (Consejo Nacional de Política Económica y Social - CONPES, 2018, p. 25).

Dato no estructurado: su organización y presentación no está guiada por ningún modelo o esquema. En esta categoría se incluyen, por ejemplo, las imágenes, textos, audios, contenidos de redes sociales, videos. (Consejo Nacional de Política Económica y Social - CONPES, 2018, p. 25).

Datos estructurados. Fuente: freepik (2059967).

Dato personal: cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables. De acuerdo con el tipo de información a la que se refieren, estos pueden ser sensibles, semiprivados, privados o públicos. Es decir, no todos los datos personales son privados (Política Nacional de Explotación de Datos (Big Data), 2018, p. 25).

Dato privado: es el dato que por su naturaleza íntima o reservada sólo es relevante para el titular. Se trata de datos que se encuentran en el ámbito propio del sujeto concernido y, por ende, sólo puede accederse por orden de autoridad judicial competente y en ejercicio de sus funciones. Entre dicha información se encuentran los libros de los comerciantes, los documentos privados, las historias clínicas, los datos obtenidos en razón a la inspección de domicilio o luego de la práctica de pruebas en procesos penales sujetas a reserva, entre otros (Ley 1581, 2012).

Dato público: es el dato que no sea semiprivado, privado o sensible. Son considerados datos públicos, entre otros, los datos relativos al estado civil de las personas, a su profesión u oficio y a su calidad de comerciante o de servidor público. Por su naturaleza, los datos públicos pueden estar contenidos, entre otros, en registros públicos, documentos públicos, gacetas y boletines oficiales y sentencias judiciales debidamente ejecutoriadas que no estén sometidas a reserva. (Decreto 1377, 2013. Artículo 3, numeral 2).

Dato semiestructurado: su organización y presentación tiene una estructura básica (etiquetas o marcadores), pero no tiene establecida una definición de relaciones en su contenido. En esta categoría se incluyen contenidos de e-mails, tweets, archivos XML. (Consejo Nacional de Política Económica y Social - CONPES, 2018).

Dato semiprivado: hace alusión al dato que no tiene naturaleza íntima, reservada, ni pública y cuyo conocimiento o divulgación puede interesar no sólo a su titular sino a cierto sector o grupo de personas o a la sociedad en general, como el dato financiero y crediticio de actividad comercial o de servicios. (Consejo Nacional de Política Económica y Social - CONPES, 2018).

Datos abiertos: son todos aquellos datos primarios o sin procesar, que se encuentran en formatos estándar e interoperables que facilitan su acceso y reutilización, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos. (Ley 1712, 2014. Artículo 6). La definición de apertura involucra los siguientes elementos: disponibilidad y acceso; reutilización y redistribución; y participación universal (OKFN, 2016). La definición de apertura implica la exclusión de todos aquellos datos que, dada su tipología, presentan limitaciones jurídicas para su acceso, divulgación, compartición y tratamiento. (Consejo Nacional de Política Económica y Social - CONPES, 2018).

Datos sensibles: aquellos que afectan la intimidad del Titular o cuyo uso indebido puede generar su discriminación, tales como aquellos que revelen el origen racial o étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, organizaciones sociales, de derechos humanos o que promueva intereses de cualquier partido político o que garanticen los derechos y garantías de partidos políticos de oposición, así como los datos relativos a la salud, a la vida sexual y los datos biométricos" (Decreto 1377, 2013).

Datos tabulados: es el resultado de la organización o agregación de los datos o microdatos, para obtener resultados definidos. (Basados en: Sanchéz, Soria-Comas, & Domingo-Ferrer, 2016, p. 4).

Documento de archivo: es el registro de información producida o recibida por una entidad pública o privada en razón de sus actividades o funciones. (Ley 1712, 2014, p. Artículo 4).

Encargado del tratamiento: persona natural o jurídica, pública o privada, que por sí misma o en asocio con otros, realice el Tratamiento de datos personales por cuenta del Responsable del Tratamiento. (Ley 1581, 2012. Artículo 3).

Gestión documental: es el conjunto de actividades administrativas y técnicas tendientes a la planificación, procesamiento, manejo y organización de la documentación producida y recibida por los sujetos obligados, desde su origen hasta su destino final, con el objeto de facilitar su utilización y conservación. (Ley 1712, 2014. Artículo 6).

Habeas data: derecho reconocido en la Constitución para conocer, actualizar y rectificar las informaciones que se hayan recogido sobre las personas en bancos de datos y en archivos de entidades públicas y privadas. Este derecho tiene una naturaleza autónoma que lo diferencia de otras garantías con las que está en permanente relación, como los derechos a la intimidad y a la información. Ver artículo 15 de la Constitución Política de Colombia (Consejo Nacional de Política Económica y Social - CONPES, 2018).

Derecho Habeas Data. Fuente: freepik (13313125).

Identificadores indirectos: son aquellos que, si bien no identifican a una persona, el cruce de varios identificadores indirectos podría permitir la identificación de una persona. (Orientaciones y garantías en los procedimientos de anonimización de datos personales, 2016, p. 12).

Información: se refiere a un conjunto organizado de datos contenido en cualquier documento que los sujetos obligados generen, obtengan, adquieran, transformen o controlen (Ley 1712, 2014. Artículo 6).

Información pública: es toda información que un sujeto obligado genere, obtenga, adquiera, o controle en su calidad de tal. (Ley 1712, 2014. Artículo 6).

Información pública clasificada: es aquella información que estando en poder o custodia de un sujeto obligado en su calidad de tal, pertenece al ámbito propio, particular y privado o semiprivado de una persona natural o jurídica por lo que su acceso podrá ser negado o exceptuado, siempre que se trate de las circunstancias legítimas y necesarias y los derechos particulares o privados. (Ley 1712, 2014. Artículo 6).

Información pública reservada: es aquella información que estando en poder o custodia de un sujeto obligado en su calidad de tal, es exceptuada de acceso a la ciudadanía por daño a intereses públicos y bajo cumplimiento de la totalidad de los requisitos consagrados en la Ley de Transparencia. (Ley 1712, 2014. Artículo 6).

Microdato: son los datos sobre las características de las unidades de una población, (individuos, hogares, establecimientos, entre otros), que constituyen una unidad de información en una base de datos y que son recogidos por medio de una operación estadística. Se refiere a un registro de información relativa a un individuo. (Ministerio de Salud y Protección Social).

Responsable del tratamiento: persona natural o jurídica, pública o privada, que por sí misma o en asocio con otros, decida sobre la base de datos y/o el Tratamiento de los datos. (Ley 1581, 2012. Artículo 3).

Seudonimización: es la sustitución de un atributo (normalmente un atributo único) por otro en un registro. Por consiguiente, sigue existiendo una alta probabilidad de identificar a la persona física de manera indirecta; en otras palabras, el uso exclusivo de la seudonimización no garantiza un conjunto de datos anónimo. (Dictamen 05 / 2014 sobre técnicas de anonimización , 10 de abril de 2014).

Titular del dato: persona natural cuyos datos personales sean objeto de Tratamiento (Ley 1581 de 2012. Artículo 3).

Tratamiento de datos: cualquier operación o conjunto de operaciones sobre datos personales, tales como la recolección, almacenamiento, uso, circulación o supresión. (Ley 1581, 2012. Artículo 3).

Transferencia de datos: la transferencia de datos tiene lugar cuando el Responsable y/o Encargado del Tratamiento de datos personales, ubicado en Colombia, envía la información o los datos personales a un receptor, que a su vez es Responsable del Tratamiento y se encuentra dentro o fuera del país. (Ley 1581, 2012. Artículo 3).

Transmisión de datos: tratamiento de datos personales que implica la comunicación de estos dentro o fuera del territorio de la República de Colombia cuando tenga por objeto la realización de un tratamiento por el encargado por cuenta del responsable. (Ley 1581, 2012. Artículo 3).

3. ANTECEDENTES EN EL USO DE LA ANONIMIZACIÓN

3.1 Contexto internacional

En su mayoría, las experiencias internacionales sobre la anonimización de información parten de indicaciones o recomendaciones de organismos propios o multilaterales. Este proceso también se ha dado como mecanismo de seguimiento a la legislación internacional en materia de protección, privacidad y confidencialidad de la información.

En el ámbito internacional dos de las legislaciones que han desarrollado el derecho a la protección de datos personales se basan en principios con enfoques diferentes. La legislación europea basada en el derecho fundamental a la privacidad y el Habeas Data, y la legislación de Estados Unidos está orientada del principio de responsabilidad.

En términos generales, la protección de datos en la Unión Europea está conformada por amplias garantías respaldadas por la constitución como derechos fundamentales integrales, y sus principios se aplican independientemente del contexto. Por el contrario, en Estados Unidos, las garantías de protección de los datos son casuísticas, específicas del contexto y de los sectores, varían según los instrumentos existentes y son mucho menos completas (European Parliament, 2015).

En el caso de la Unión Europea la compartición de datos entre agencias compromete los derechos fundamentales de las personas y requieren de una justificación concreta, mientras que en Estados Unidos el intercambio de datos sin mayor restricción es más la regla que la excepción (European Parliament, 2015). Otra importante distinción entre ambos marcos se determina por el alcance de la ley de protección de datos personales. En la Unión Europea la posible vulneración de derechos fundamentales a partir del tratamiento de datos brinda la posibilidad inmediata para que el individuo inicie un proceso judicial, mientras que, en Estados Unidos, la recopilación masiva de datos no conduce inmediatamente a otorgar una acción jurídica al sujeto titular de sus datos.

Por otra parte, en el Reino Unido, la Oficina del Comisionado de Información desarrolló el Código de Buenas Prácticas para la Anonimización (Oficina del Comisionado de información, 2012), presentando los antecedentes jurídicos para la protección de datos de ese país; explica los beneficios de la anonimización, todo enmarcado en los principios que deben guiar dicho proceso. Se señalan también los riesgos que se pueden materializar al cruzar información en bases de datos que ya se encuentran anonimizadas, generando posibles identificaciones de usuarios y el control de los datos en los diferentes sistemas de información.

3.2 Contexto nacional

Colombia ha gestionado la implementación de un marco legal en esta materia, partiendo del derecho a la intimidad personal y familiar y el derecho a conocer, actualizar y rectificar las informaciones que se hayan recogido en las bases de datos y en archivos de entidades públicas y privadas, que aparece consignado en el artículo 15 de la Constitución Política de 1991.

En términos de la confidencialidad, enmarcados en las directrices para el Sistema Estadístico Nacional - SEN, es importante tener en cuenta la Ley Estatutaria 1266 de 2008 que establece el Habeas Data y regula el manejo de la información contenida en bases de datos personales, financieras, crediticias, comerciales, de servicios y provenientes de terceros países, además de establecer disposiciones sobre la recolección, tratamiento y circulación de datos personales en el país (Congreso de Colombia, 2008).

Además de estas leyes, se cuenta con la Ley 1581 de 2012, que trata sobre la protección de datos personales, reglamentada parcialmente por el decreto 1377 de 2013, en la cual se establece el acceso a los datos se debe restringir y la información debe estar sujeta a tratamiento por parte del responsable, como lo indica en su artículo 4, manteniendo los principios de acceso y circulación restringida, de seguridad y de confidencialidad.

Ley 1581 de 2012. Fuente: freepik (20916445).

Así mismo se cuenta con la Ley 79 de 1993, en la cual se establece mantener la confidencialidad de las fuentes cuando se realizan procesos de recolección de información a través de censos o encuestas.

Respecto a la transparencia y el acceso de la información pública, la Ley 1712 de 2014, regula el derecho de acceso a la información pública, haciendo énfasis en el establecimiento de una política de datos abiertos por parte de las entidades públicas. Además, el Decreto 2573 de 2014 del Ministerio de Tecnología de la información y las Comunicaciones (MinTIC), establece los Lineamientos Generales de la Estrategia de Gobierno en línea, indicando los principios y fundamentos a tener en cuenta en las entidades públicas destacando entre ellos, la excelencia en el servicio ciudadano, la apertura y reutilización de datos públicos, la estandarización, la innovación, entre otros. Igualmente, se establecen cuatro componentes que facilitarán la masificación de la oferta y la demanda en gobierno en línea, resaltando entre estos, la seguridad y la privacidad de la información, siendo un componente transversal.

Por otro lado, el CNMH cuenta con la Política de Seguridad y Privacidad de la Información en su versión 2 y código SIP-PC-013, la cual es aplicable a todos los procesos y aspectos administrativos y de control que deben ser cumplidos por los funcionarios, contratistas y terceros que presten sus servicios o tengan algún tipo de relación con el Centro Nacional de Memoria Histórica - CNMH, para el cumplimiento de sus funciones y para obtener un adecuado nivel de protección de la seguridad de la información, participando en la toma de medidas preventivas y correctivas. Esta política está contenida en el Sistema de Gestión de Seguridad de la Información – SGSI y es complementaria al presente documento.

Archivos. Fuente: freepik (21311212).

4. OBJETIVO Y ALCANCE DE LA GUÍA

Presentar estándares, lineamientos técnicos y una orientación metodológica para realizar procesos de anonimización de datos personales, información y documentos textuales, con el fin de salvaguardar datos sensibles durante el procesamiento técnico y puesta el servicio en el Archivo virtual.

La guía está dirigida a funcionarios y servidores públicos, grupos o áreas de las entidades públicas y privadas con funciones públicas e instancias y nuevas instituciones en el marco de la justicia transicional, que gestionen, almacenen, administren, obtengan, produzcan, procesen, custodien y publiquen información, independientemente de su soporte o medio, y que requieran la técnica de anonimización.

5. MARCO CONCEPTUAL

5.1 Anonimización

Es un proceso técnico que hace que la información de identificación sobre las personas no sea identificable a quienes pertenece la información. Anonimizar permite ocultar la información que identifique plena o parcialmente a las personas, organizaciones o características individuales de la fuente de información.

La anonimización se realiza para facilitar la divulgación, la publicación y el intercambio de datos, sin vulnerar los derechos a la protección de datos, de manera que no se puedan identificar directa o indirectamente las personas asociadas a dicha información de identificación personal.

La finalidad del proceso de anonimización es eliminar o reducir al mínimo los riesgos de reidentificación de los datos anonimizados manteniendo la veracidad de los resultados del tratamiento de estos, es decir, además de evitar la identificación de las personas, los datos anonimizados deben garantizar que cualquier operación o tratamiento que pueda ser realizado con posterioridad a la anonimización no conlleva una distorsión de los datos reales. Un análisis masivo de los datos que puedan derivar de los datos anonimizados no debería diferir del análisis que pudiera obtenerse si hubiera sido realizado con datos no anonimizados.

En el proceso de anonimización se deberá producir la ruptura de la cadena de identificación de las personas. Esta cadena se compone de microdatos o datos de identificación directa y de datos de identificación indirecta[4]. Los microdatos

permiten la identificación directa de las personas y los datos de identificación indirecta son datos cruzados de la misma o de diferentes fuentes que pueden permitir la reidentificación de las personas, como la información de otras bases de datos del mismo u otro responsable, de las redes sociales, buscadores, blogs, etc.

En el diseño del proceso de anonimización será necesario prever las consecuencias de una eventual reidentificación de las personas que pudiera generar un perjuicio o merma de sus derechos. Igualmente será necesario prever una hipotética pérdida de información por negligencia de personal implicado, por falta de una política de anonimización adecuada o por una revelación de secreto intencionada que diera lugar a la pérdida de las variables de identificación o claves de identificación de las personas.

5.2 Principios

La guía adopta los principios constitucionales en general, los de la Ley 1581 de 2012 y los consignados en el CONPES 3920 de 2018, expresados de la siguiente manera:

Respeto de los derechos humanos: las entidades y las personas encargadas de los procesos de tratamiento de datos personales y anonimización respetarán los derechos fundamentales reconocidos en la Constitución y en los instrumentos internacionales que los consagran. Todas sus actuaciones contribuirán a la garantía de los Derechos Humanos y el Derecho internacional Humanitario, entendiéndose estos como el límite de la explotación de datos y por tanto guiando las decisiones de la anonimización.

Legalidad: la anonimización debe cumplir con la protección de datos personales según la normatividad vigente y debe tener siempre fines lícitos.

Finalidad: la anonimización debe cumplir la normatividad en materia de protección de datos personales y debe tener siempre fines lícitos.

Acceso restringido: la anonimización debe garantizar que los datos personales no se divulgarán, que si fuera necesario se supriman o rectifiquen datos que permitan la re-identificación. Las entidades y personas encargadas de la anonimización están obligadas a garantizar la clasificación de la información, incluso luego de la anonimización. Las entidades y personas responsables de la anonimización deben responder al cumplimiento de la legislación vigente por medio de la rendición de cuentas y con la respuesta a solicitudes de los órganos de control sobre las formas, técnicas y decisiones sobre la anonimización de datos.

Seguridad: se debe garantizar la seguridad de los datos personales y el resultado de la anonimización para evitar el tratamiento no autorizado, ilegal o adulteración y su pérdida (por destrucción o daño accidental). Se deben tomar todas las medidas técnicas, humanas y administrativas necesarias.

Utilidad: se debe tener en cuenta la utilidad final de los datos anonimizados y garantizar la inexistencia de alteraciones que impidan utilizar la información anonimizada de acuerdo con la finalidad que se pretenda.

Proactividad: se debe garantizar la confidencialidad de la información personal desde el diseño o puesta en marcha del proceso de anonimización mediante la técnica seleccionada.

6. ¿SE DEBE ANONIMIZAR?

La anonimización de datos, al eliminar las posibilidades de identificación de las personas u organizaciones, lo que busca principalmente es proteger la identidad y los datos asociados a esa identidad.

Se debe anonimizar para proteger los derechos de los titulares de los datos e información y reducir o eliminar definitivamente el riesgo de reidentificación. Pero debe permitir el uso de la información de manera adecuada para los fines establecidos.

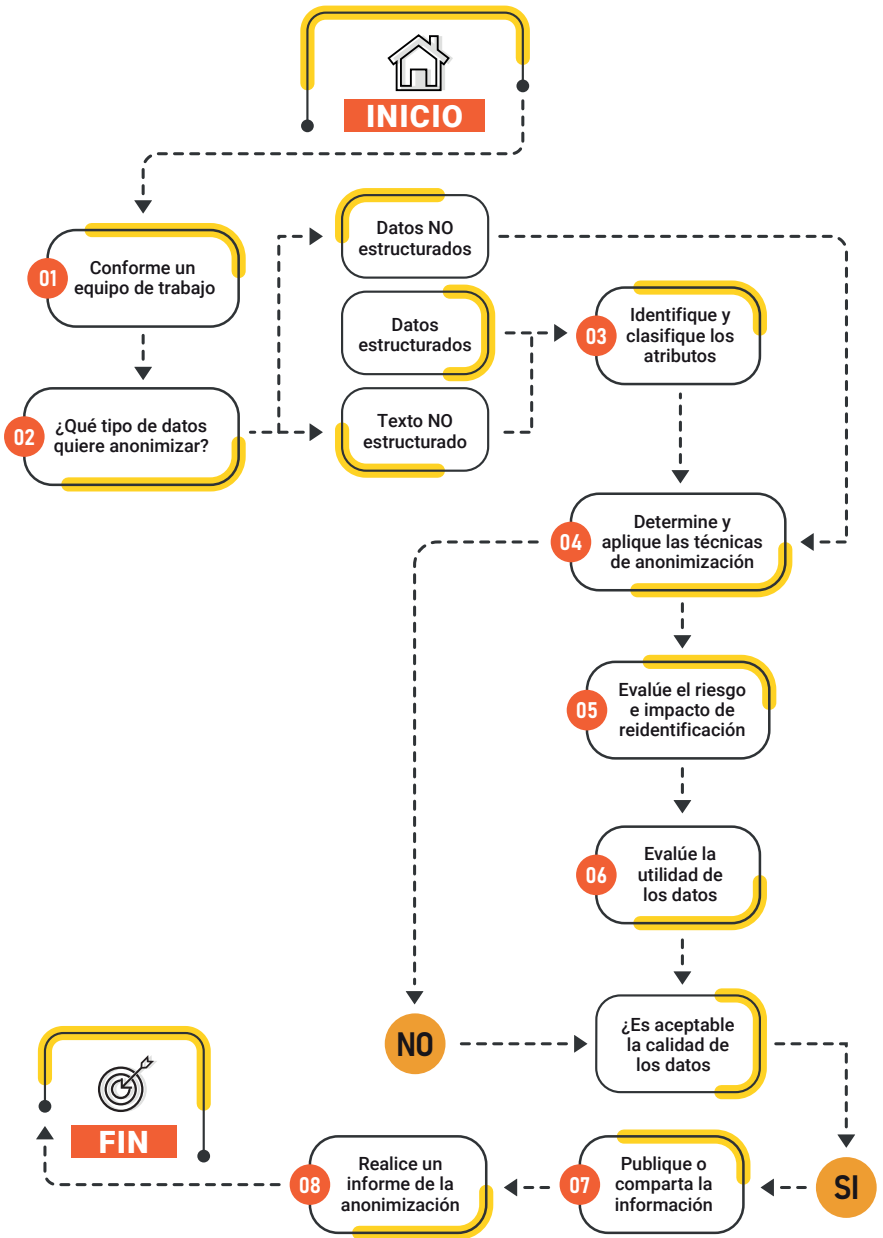
Se debe anonimizar para evitar no solo la identificación directa, sino también la identificación indirecta: esta se obtiene del cruce de datos y otras fuentes de información.

La anonimización de los datos e información debe prever los riesgos de reidentificación y la pérdida de datos.

Se debe aplicar cuando se va a publicar información y no se requiere de información específica.

Se debe aplicar cuando se intercambia información con otras entidades y este intercambio no implica la necesidad de identificación o la inclusión de datos personales.

Figura 1. Diagrama para la anonimización de datos.



7. METODOLOGÍA: ¿QUÉ PASOS SE DEBEN SEGUIR PARA ANONIMIZAR?

A continuación, se propone un flujo de actividades que permiten orientar a las entidades como definirlos; en ningún caso se trata de un modelo cerrado, sino que es una estructura a tener en cuenta en los procesos de anonimización que se pretendan realizar. En ningún caso se trata de un modelo cerrado, sino que es una estructura a tener en cuenta en los procesos de anonimización que se pretendan realizar.

7.1 Conformar un equipo de trabajo

La entidad debe conformar un equipo de trabajo, liderado por la Dirección de Archivo de los Derechos Humanos y el área de seguridad informática, de la dirección de tecnologías de la información o la que haga sus veces. Se recomienda que a este grupo se vincule también el área de gestión documental o la que haga sus veces. Los roles pueden ser desempeñados por una o varias personas y se recomienda que sean tenidos en cuenta los siguientes:

- Director Técnico de la Dirección de Archivo de Derechos Humanos.
- Director de sistemas y tecnologías de la Información o el que haga sus veces.
- Director de gestión documental o el que haga sus veces.
- Encargado de la seguridad de la información de la entidad o el que haga sus veces.
- Asesores o profesionales expertos en temas específicos de los que trate la información o los datos a anonimizar (expertos temáticos).
- Profesional de anonimización. Se recomienda que en lo posible los profesionales de anonimización deben ser profesionales que no estén en contacto directo o administren las bases de datos o sistemas de información en la entidad. Oficial de protección de datos personales.

7.2 Identificar qué tipo de datos requiere anonimizar

Identificar los datos que son estructurados, es decir, aquellos que están organizados en forma de tablas o matrices que pueden estar en formato de Excel o en bases de datos que contienen los microdatos.

Igualmente se debe identificar los datos no estructurados, o que no corresponden a ningún modelo establecido, estos pueden ser: imágenes, textos, audios, contenidos de redes sociales, audiovisuales.

7.3 Identificar y clasificar los atributos

Los atributos de los microdatos pueden ser:

- Identificadores directos: son todas aquellas características que por sí mismas permiten la identificación de una persona u entidad de manera inequívoca dentro de un conjunto de datos.
- Identificadores indirectos o cuasi - identificadores: son aquellas características que por sí solas no permiten la identificación de una persona u entidad, pero que relacionados o en combinación con otros identificadores indirectos podrían permitir la identificación dentro de un conjunto de datos.

Se debe tener en cuenta que existen identificadores directos e indirectos de tipo clasificado[5] que consisten en características de datos que pueden afectar la intimidad del titular y su uso indebido puede generar violaciones a derechos, tales como el origen étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos u organizaciones sociales, y los relativos a la salud, la vida sexual y los datos biométricos.

7.4 Determinar y aplicar las técnicas de anonimización

Las técnicas de anonimización a implementar dependen del tipo de identificadores determinados en el paso 7.3 y del uso que se le va a dar a los datos.

La implementación de las técnicas de anonimización debe ser evaluada y aprobada por el equipo de trabajo.

La aplicación de las técnicas de anonimización puede ser realizada manualmente (por ejemplo en hojas de cálculo) o se pueden usar administradores o herramientas de bases de datos. Para la selección de herramientas se debe tener en cuenta lo siguiente:

- Que puedan soportar la aplicación de diferentes técnicas de anonimización.
- Que permitan la anonimización para datos estructurados y no estructurados.
- Que permitan el cálculo del riesgo de reidentificación.
- Que permitan un fácil acceso a evaluar la calidad de la anonimización.

7.5 Evaluar el riesgo e impacto de reidentificación

Las técnicas de anonimización aplicadas, deben tener definidas las posibilidades de reidentificación. Se deben tener en cuenta las condiciones de gestión de la información para evitar riesgos:

- Se debe tener claridad acerca de la finalidad de la publicación o del intercambio de información.
- La entidad debe tener claramente señalados los niveles de acceso, las responsabilidades y la circulación interna y externa de la información sin anonimizar y de la anonimizada.
- La entidad debe garantizar la formación y la información con la que cuenta el equipo de trabajo encargado de los procedimientos.
- Siempre se debe calcular el riesgo teniendo en cuenta que puede existir interés de identificar a un individuo dentro de un conjunto de datos anonimizados ya publicado.
- Se debe calcular el riesgo independientemente del tipo de datos anonimizados que se pública y de que se identifique un posible interés en reidentificar.

Usando cálculos y modelos de probabilidad pueden establecerse los riesgos y las probabilidades de ocurrencia de estos, probabilidades que deben partir de los factores de daño que se causarían a los titulares en caso de reidentificación. También se debe calcular el daño que causaría a la entidad la reidentificación de las personas.

Se pueden usar categorías como: riesgos de reidentificación existentes y conocidos, riesgos potenciales y riesgos no conocidos. Los cálculos y modelos usados deben ser coherentes y adecuados de acuerdo con las características y el volumen de la información que se procesa.

Una vez identificado el nivel de riesgo de reidentificación en los resultados finales del proceso de anonimización, se deben evaluar periódicamente los impactos de la reidentificación.

Se deben tener en cuenta los siguientes aspectos dentro de la evaluación de los riesgos:

- Riesgos originados en la aplicación inadecuada de los procedimientos y las técnicas.
- El riesgo se puede incrementar al agregar más datos a lo largo del tiempo o como resultado del desarrollo de programas o proyectos específicos (por ejemplo, la atención a víctimas, la ampliación del cubrimiento en educación, etc.).
- El riesgo se incrementa en cuanto a los identificadores indirectos, por parte del titular de la información por medio del uso de redes sociales y otros medios de comunicación.

Las revisiones periódicas tienen por finalidad verificar que el estado real de riesgos coincide con el riesgo asumible de reidentificación, verificando la eficacia de las medidas previstas para paliar el posible impacto que pudiera tener la reidentificación de las personas. (Agencia Española para la protección de datos, 2016)

Los riesgos que pueden presentarse durante el uso de procedimientos de anonimización por parte de las entidades públicas, son los siguientes:

Singularización: *se trata de la posibilidad de extraer de un conjunto de datos algunos registros (o todos los registros) que identifican a una persona.*

Este riesgo implicaría que los datos todavía presentan rasgos que permiten identificar a las personas, ya que todavía no han perdido su condición de datos semiprivados, privados o sensibles y por eso pueden servir para particularizar al individuo.

Vinculabilidad: *la capacidad de vincular o combinar como mínimo dos registros de un único interesado o de un grupo de interesados, ya sea en la misma base de datos o en dos bases de datos distintas. Si el atacante puede determinar (p. ej., mediante un análisis de correlación) que dos registros están asignados al mismo grupo de personas, pero no puede singularizar a las personas en este grupo, entonces la técnica es resistente a la singularización, pero no a la vinculabilidad.*

Inferencia: *la posibilidad de deducir con una probabilidad significativa el valor de un atributo a partir de los valores de un conjunto de otros atributos.*

Ejemplo de identificación digital por huella. Fuente: freepik (5330027).

7.6 Evaluar la utilidad de los datos

El proceso de anonimización debe realizarse bajo la premisa general de que los datos publicados o intercambiados de esta manera, pueden perder características que disminuyen su utilidad. Se tiene que evaluar esta situación en el marco del equilibrio entre el nivel de privacidad o la protección de los datos personales y la utilidad.

Esta utilidad puede ser evaluada por el grupo de trabajo y se relaciona directamente con que su utilidad sea medida dependiendo del uso que se le va a dar después de anonimizar los datos. Esto puede resultar complejo cuando los usuarios potenciales pueden ser muchos y no se pueden determinar todos los usos que se le van a dar a los datos.

7.7 Publicar o compartir la información

Si los datos a anonimizar son un conjunto de datos abiertos con destino al portal de la entidad o es para intercambiar con otra entidad, se recomienda que estos procesos sean automatizados para facilitar su publicación o intercambio haciendo uso de las herramientas del Manual de Gobierno Digital. Allí se encuentra el marco de interoperabilidad del Estado y las guías para publicación de datos abiertos dispuestas por el Ministerio de Tecnologías de la Información y las Comunicaciones.

En cualquier caso, para cualquier publicación e intercambio se recomienda contar con medidas de seguridad de la información adecuadas, para garantizar que en esa publicación o intercambio de datos no se pierda, altere, sustraiga o modifique la información y se afecten los derechos de las personas. Adicionalmente se debe documentar la trazabilidad de la gestión.

Así mismo, una vez anonimizados los datos, se deben actualizar los inventarios de activos de información de la entidad y si son necesarios los índices de información clasificada o reservada, dado que estos se configuran como un conjunto de datos diferente y es considerado otro activo de información.

7.8 Realizar un informe de anonimización

El informe de anonimización debe realizarse detallando los procedimientos, técnicas y metodologías usadas, debe contener la trazabilidad de las gestiones desde el tratamiento de los datos iniciales hasta obtener los datos anonimizados. Además, debe incluir el análisis de riesgos con sus respectivos cálculos y anexos.

Este informe debe reflejar, todos los aspectos necesarios para evaluar la adecuada anonimización posteriormente y debe contener como mínimo:

- Responsables del proceso de anonimización y equipo que participó en el proceso.
- Fecha o intervalo de tiempo que duró el proceso de anonimización.
- Técnicas de anonimización empleadas.
- Riesgos e impactos analizados.

Este informe debe estar incluido en la tabla de retención documental (TRD) de la dependencia que asume la responsabilidad del manejo de la información o a la que pertenezca el grupo de trabajo que realiza los procesos de acuerdo con la estructura orgánico funcional de la Entidad. En todo caso, también debe quedar en el repositorio de evidencias de los procesos de la entidad.

Ejemplos varios de procesos de anonimización. Fuente: freepik (adaptación imagen 22551576).

8. TÉCNICAS DE ANONIMIZACIÓN PARA DATOS ESTRUCTURADOS

8.1 Aplicación de técnicas de anonimización

En esta fase el grupo de trabajo aplica las técnicas de anonimización adecuadas y necesarias que aseguren que no sea posible identificar o reidentificar a las personas naturales o jurídicas que suministraron los datos personales a los organismos públicos o a las entidades privadas que cumplan funciones públicas.

En esta etapa es vital que la información recolectada cuente con los criterios de seguridad y confidencialidad y las medidas de seguridad necesarias tanto internas como externas.

8.1.1 Métodos de aleatorización o perturbación

Se refieren a procedimientos que implican la modificación sistemática de datos (a veces en pequeñas cantidades aleatorias), de manera tal que las cifras no sean lo suficientemente precisas como para revelar información sobre casos individuales. Pueden incluirse nuevos datos, suprimir y/o modificar los existentes beneficiando la confidencialidad estadística. (Departamento Administrativo Nacional de Estadística, 2014, p. 19).

Son métodos que modifican la veracidad de los datos para eliminar el vínculo entre los mismos y el individuo. Si los datos se hacen lo suficientemente ambiguos, no podrán remitir a una persona concreta. La aleatorización por sí sola no reduce la singularidad de cada uno de los registros, ya que estos pueden obtenerse a partir de un único interesado, pero sí puede proteger contra ataques o riesgos de inferencia. (Grupo de trabajo sobre protección de datos. Comisión Europea, 10 de abril de 2014, p. 14).

Los principales métodos de aleatorización o perturbación de datos son los siguientes:

Microagregación: la microagregación se aplica cuando las variables son numéricas y reemplazan un conjunto integrado por K datos con la media calculada sobre ese mismo conjunto. El K mínimo aceptado es 3, sin embargo, en la elección del K se debe sopesar la pérdida de información y de variabilidad. (Ministerio de Salud y Protección Social).

La idea es reemplazar un valor observado con la media calculada sobre un pequeño grupo de unidades (agregado pequeño o micro-agregado), incluido el

investigado. Consiste en agrupar los registros individuales en pequeños grupos antes de su publicación, manteniendo los resultados al aplicar las operaciones estadísticas. (Departamento Administrativo Nacional de Estadística, 2017).

Adición de ruido: consiste en modificar los atributos del conjunto de datos para que sean menos exactos, conservando no obstante su distribución general. (Grupo de trabajo sobre protección de datos. Comisión Europea, 10 de abril de 2014, p. 13).

Permutación o intercambio de registros: esta permutación garantiza que el rango y la distribución de valores sean idénticos además de que mantiene el nivel de detalle; no obstante, las correlaciones entre los valores y los individuos pueden quedar destruidas. (Ministerio de Salud y Protección Social).

Redondeo: consiste en la sustitución del valor de las variables originales por valores redondeados de forma aleatoria. La variable debe ser numérica. (Departamento Administrativo Nacional de Estadística, 2014, p. 22).

Reajuste de peso: en los casos donde se conoce el tipo de muestreo utilizado es posible revertir el proceso, lo cual aumenta la posibilidad de identificar de manera puntual a una persona; por tanto, es conveniente hacer una modificación de los pesos con el fin de disminuir este riesgo. (Ministerio de Salud y Protección Social).

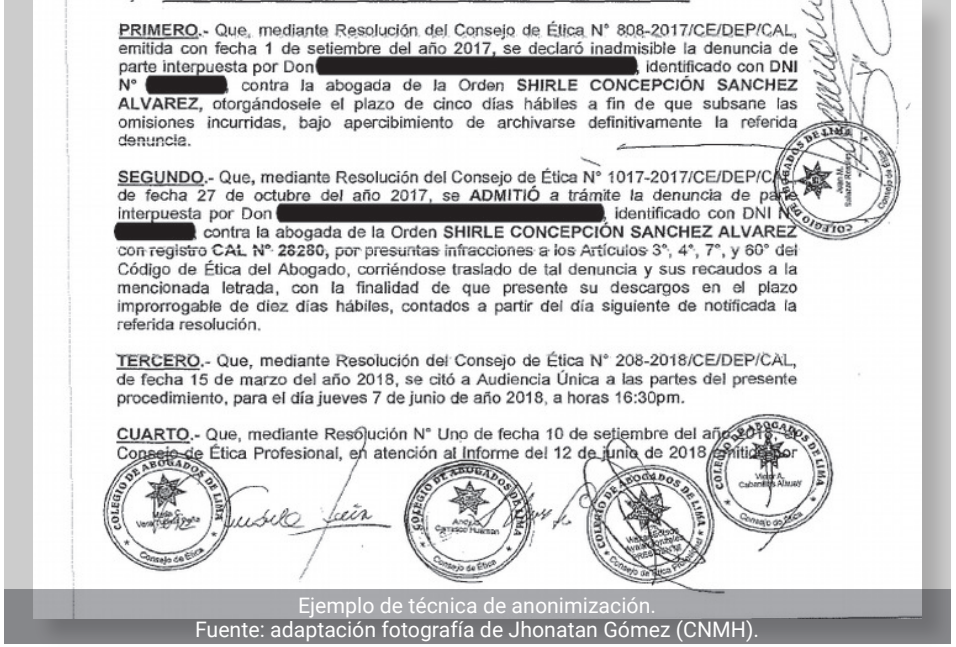
8.1.2 Métodos de reducción o generalización

Existen métodos basados en la reducción de datos en donde aplicando estas técnicas no se alteran los datos, sino que producen supresiones parciales o reducciones del nivel de detalle del conjunto original. Estos procedimientos tienden a evitar la presencia de individuos reconocibles únicos o atípicos. (Departamento Administrativo Nacional de Estadística, 2014, p. 23).

Son métodos que generalizan o diluyen los atributos de los interesados modificando las respectivas escalas u órdenes de magnitud (por ejemplo, sustituyendo una ciudad por una región, o una semana por un mes). (Grupo de trabajo sobre protección de datos. Comisión Europea, 10 de abril de 2014, p. 18).

Las principales técnicas de reducción o generalización de datos son las siguientes:

Eliminación de variables: la primera aplicación de este método es la eliminación de identificadores directos desde el archivo de datos. Una variable debe eliminarse cuando está muy identificada y no puede aplicarse otro método de


PRIMERO.- Que, mediante Resolución del Consejo de Ética N° 808-2017/CE/DEP/CAL, emitida con fecha 1 de setiembre del año 2017, se declaró inadmisible la denuncia de parte interpuesta por Don [redacted], identificado con DNI N° [redacted], contra la abogada de la Orden SHIRLE CONCEPCIÓN SANCHEZ ALVAREZ, otorgándosele el plazo de cinco días hábiles a fin de que subsane las omisiones incurridas, bajo apercibimiento de archivarse definitivamente la referida denuncia.

SEGUNDO.- Que, mediante Resolución del Consejo de Ética N° 1017-2017/CE/DEP/CAL de fecha 27 de octubre del año 2017, se ADMITIÓ a trámite la denuncia de parte interpuesta por Don [redacted], identificado con DNI N° [redacted], contra la abogada de la Orden SHIRLE CONCEPCIÓN SANCHEZ ALVAREZ con registro CAL N° 26280, por presuntas infracciones a los Artículos 3°, 4°, 7°, y 60° del Código de Ética del Abogado, corriéndose traslado de tal denuncia y sus recaudos a la mencionada letrada, con la finalidad de que presente su descargos en el plazo improrrogable de diez días hábiles, contados a partir del día siguiente de notificada la referida resolución.

TERCERO.- Que, mediante Resolución del Consejo de Ética N° 208-2018/CE/DEP/CAL, de fecha 15 de marzo del año 2018, se citó a Audiencia Única a las partes del presente procedimiento, para el día jueves 7 de junio de año 2018, a horas 16:30pm.

CUARTO.- Que, mediante Resolución N° Uno de fecha 10 de setiembre del año 2018, Consejo de Ética Profesional, en atención al Informe del 12 de junio de 2018 emitido por

Ejemplo de técnica de anonimización.
Fuente: adaptación fotografía de Jhonatan Gómez (CNMH).

protección. También puede quitar una variable cuando es demasiado sensible para uso público o irrelevante a efectos analíticos. (Departamento Administrativo Nacional de Estadística, 2014, p. 24)

Eliminación de registros: puede adoptarse como medida extrema de protección de datos, cuando la unidad es identificable a pesar de la aplicación de otras técnicas de protección. (Departamento Administrativo Nacional de Estadística, 2014, p. 24).

Recodificación global: cuando se agrupan determinadas categorías de datos en una nueva categoría reduciendo las posibilidades de reidentificación. (Agencia Española para la protección de datos, 2016, p. 17).

Supresión de celdas: si al hacer cruces de información se muestran celdas que pueden revelar información (frecuencias bajas) que conduzcan a la identificación de individuos se puede suprimir una o varias celdas de la tabla. Esta técnica es útil también cuando se trata de información presentada en gráficos. (Ministerio de Salud y Protección Social).

8.1.3 Método o técnica de preanonimización

En el momento en que se esté concibiendo el diseño de la operación estadística, se deberá realizar la etapa de preanonimización, entendida como un paso previo que permite determinar con claridad de las variables, los identificadores directos y demás datos con carácter confidencial que se obtendrán en el desarrollo de la operación estadística. (Departamento Administrativo Nacional de Estadística, 2014, p. 11).

8.1.4 Método o técnica de seudoanonimización

Es un método utilizado para ocultar identidades. La finalidad del uso de seudónimos es poder recopilar más datos sobre una misma persona sin necesidad de conocer su identidad. Su uso es especialmente pertinente en los ámbitos estadísticos e investigativos y deben evaluarse los riesgos de identificación directa. (Departamento Administrativo Nacional de Estadística, 2014, p. 8).

El DANE (Lineamientos para la anonimización de microdatos, 2014, p. 14) sugiere los siguientes pasos:

- **a.** Asignar un único seudónimo a cada objeto de la información personal identificable.
- **b.** El seudónimo debe ser utilizado en reemplazo de números de identificación como cédula, licencias de conducción, etc. Se recomienda que los seudónimos tengan la misma longitud y formato para aumentar la legibilidad.
- **c.** Tener en cuenta el impacto de los sistemas de información en la asignación de los seudónimos en relación con los usos internos.
- **d.** Si se utilizan seudónimos para uso externo, estos deben ser diferentes a los seudónimos generados para uso interno, y no tener una relación entre uno y otro.
- **e.** El equipo de sistemas deberá establecer las técnicas criptográficas para llevar a cabo la incorporación de seudónimos que reemplacen las variables de identificación directa.

Figura 2. Ejemplo de seudoanonimización.

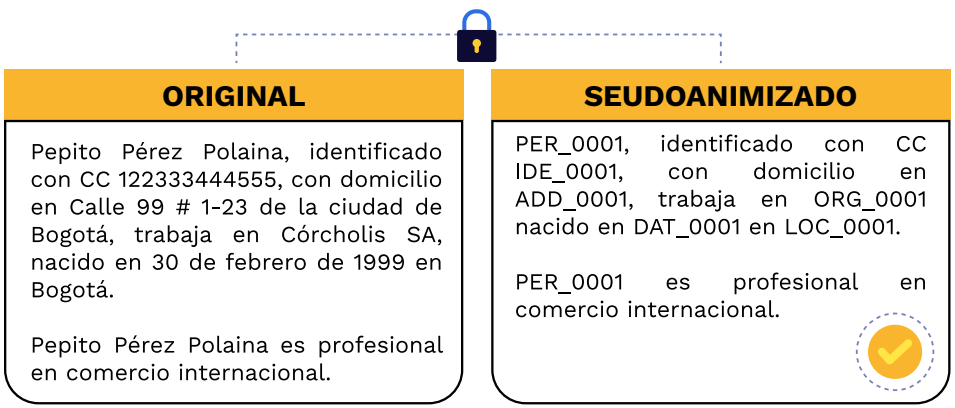


Tabla 1. Listado de atributos o identificadores vs técnica de anonimización

IDENTIFICADOR	DIRECTO	INDIRECTO	TÉCNICAS DE ANONIMIZACIÓN
Número de identificación personal	x		Eliminación de variables Supresión de celdas
Nombre completo	x		Eliminación de variables Supresión de celdas Recodificación global
Correo electrónico	x		Eliminación de variables Supresión de celdas
Número de pasaporte	x		Eliminación de variables Supresión de celdas
Número de teléfono		x	Eliminación de variables Supresión de celdas
Código postal		x	Recodificación global
Barrio/Localidad		x	Recodificación global
Ciudad o municipio de residencia		x	Recodificación global
Región		x	Recodificación global
Archivo de voz	x		Eliminación de variables Supresión de celdas
Fotografía personal	x		Eliminación de variables Supresión de celdas
Fecha de nacimiento		x	Recodificación global
Edad		x	Recodificación global
Género		x	Recodificación global
Estado civil		x	Recodificación global
Composición del hogar		x	Recodificación global
Ocupación		x	Recodificación global
Sector donde labora		x	Recodificación global
Estatus de empleo (desempleado - empleado - independiente)		x	Recodificación global
Nivel de escolaridad		x	Recodificación global
Profesión		x	Recodificación global
Idioma		x	Recodificación global
Nacionalidad		x	Recodificación global
Entidad donde labora		x	Recodificación global
Placa del vehículo	x		Eliminación de variables Supresión de celdas

IDENTIFICADOR	DIRECTO	INDIRECTO	TÉCNICAS DE ANONIMIZACIÓN
Página web	x		Eliminación de variables Supresión de celdas
Código estudiantil	x		Eliminación de variables Supresión de celdas
Número de licencia de conducción	x		Eliminación de variables Supresión de celdas
Número de cuenta bancaria	x		Eliminación de variables Supresión de celdas
Dirección IP		x	Recodificación global
Grupo étnico		x	Recodificación global
Religión		x	Recodificación global
Orientación sexual		x	Recodificación global
Número de historia clínica	x		Eliminación de variables Supresión de celdas

9. TÉCNICAS DE ANONIMIZACIÓN PARA DOCUMENTOS FÍSICOS

9.1 Documentos textuales en soporte papel y en soportes analógicos

Para los documentos que se encuentran en soporte papel y se requiere de la anonimización de alguno de sus contenidos o datos, deben tenerse en cuenta los mismos principios y metodologías contenidas en esta guía y además lo siguiente:

- Los documentos en cualquier soporte y formato deben cumplir con la organización archivística, es decir, deben estar incluidos y aplicárseles todas las herramientas y procesos archivísticos.
- Para procesos de digitalización y anonimización de documentos en soporte papel se deben realizar los estudios técnicos adecuados, teniendo en cuenta los aspectos de: conservación física, condiciones ambientales, condiciones operacionales, la seguridad de la información y la preservación a largo plazo.
- Los documentos reproducidos en medios diferentes a los usados para su producción "gozarán de la validez y eficacia del documento original, siempre que se cumplan los requisitos exigidos por las leyes procesales y se garantice la autenticidad, integridad e inalterabilidad de la información".
- "Los documentos originales que posean valores históricos no podrán ser destruidos, aun cuando hayan sido reproducidos y/o almacenados mediante cualquier medio".
- La divulgación parcial de los contenidos de los documentos, que no estén protegidos por las excepciones de la Ley 1712 de 2014, puede hacerse mediante la producción de una versión pública proveniente y certificada por la misma entidad productora.

Para los documentos físicos que contengan información y datos susceptibles de anonimizar se sugiere tener en cuenta:

- El acceso a los documentos debe controlarse para las personas no autorizadas.
- El acceso se debe limitar a los funcionarios autorizados para su gestión y esta autorización debe ser normalizada y reglada de acuerdo con los lineamientos de la entidad.
- Los documentos clasificados y con reserva, según los términos de la Ley 1712 de 2014 deben tener actualizados sus respectivos índices y los protocolos y medidas de seguridad.

Los expedientes en soporte papel pueden necesitar anonimizarse de manera total o parcial. Para retirar los documentos de un expediente debe indicarse, el lugar en donde se encuentran por medio de un testigo (formato en papel para tal fin). Si son varios documentos el testigo puede incluirse al principio del expediente incluyendo un listado de lo que se retira.

Para anonimizar, se hace una copia (fotocopia) del documento original y de esta copia se retira la información que se requiere (copia 1). Se hace una copia de esta (copia 2), y es esta la que se dará a conocer en forma de fotocopia o digitalizada. Se sugiere marcar las copias 1 y 2 para identificarlas a partir del original. Al realizar este procedimiento también debe realizarse un informe sobre las acciones llevadas a cabo, las razones de la anonimización y las partes retiradas, incluyendo la copia 1 se deben eliminar de manera adecuada. El documento original se deberá conservar sin afectar su integridad.

Es posible que en la copia 1 se retire la información por medio un marcador grueso de tinta y se saque una copia 2 que se revisará para ser divulgada evitando que queden trazos o señales de la información retirada.

Si la información a anonimizar son párrafos completos es posible colocar un papel y así emitir una copia 1 y de esta una copia 2 que será la que se divulga.

Se recomienda que estos procesos no se realicen en lugares de acceso público o en máquinas fuera de las entidades para evitar que se pueda restaurar la información.

Se recomienda que siempre se compruebe la copia 2 para asegurarse que no se puede distinguir ninguna información anonimizada.

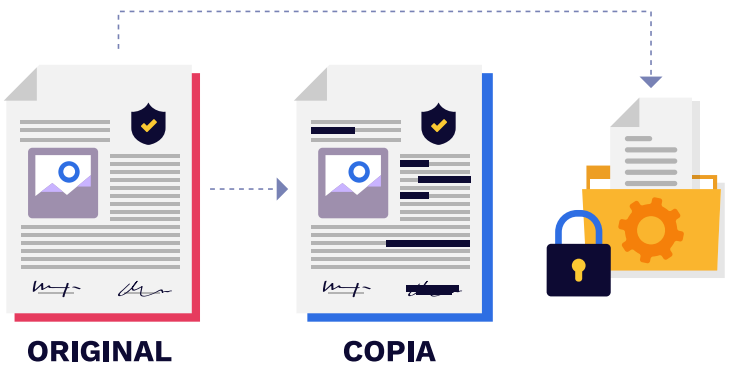
Se debe guardar la copia 2, que constituye el documento anonimizado, conservando los códigos y la signatura topográfica original (fondo, serie, subserie, caja, carpeta) en un expediente nuevo para asegurar la trazabilidad del procedimiento y garantizar el acceso a esta información si es requerida nuevamente.

9.2 Documentos textuales en formatos digitales (texto no estructurado)

Para la anonimización de estos documentos se sugiere usar una metodología análoga a la descrita para los documentos en papel, de manera que se preserve el original y se expida una copia que contenga el borrado de las partes restringidas que será pública y guardando esta copia anonimizada.

Para esto pueden usarse recursos disponibles de intervención de los archivos electrónicos como editores y programas de diseño gráfico.

Figura 3. Ejemplo: anonimización de documento físico.



10. TÉCNICAS DE ANONIMIZACIÓN DE DATOS NO ESTRUCTURADOS DE TIPO AUDIOVISUAL

El lenguaje audiovisual contiene tipos de información abundante en datos cualitativos con usos muy variados, que para su creación implican una diversidad de procesos y técnicas con características particulares, dependiendo el género (cine, música, video, etc.), el formato y el soporte. Para acotar los procesos de anonimización de este tipo de datos, es importante tener presente estas particularidades. De esta forma,

> La desidentificación en contenido multimedia se define como el proceso [reversible] de ocultar o eliminar identificadores personales, o sustituirlos por identificadores personales sustitutos en contenido multimedia, a fin de evitar la divulgación y el uso de datos para fines no relacionados con el propósito para el que la información era originalmente obtenida… La anonimización se refiere al proceso de desidentificación de datos, que produce datos donde los registros individuales no pueden vincularse con un original, ya que no incluyen las variables de traducción requeridas para hacerlo. Es un proceso unidireccional (irreversible) y no permite que los identificadores originales se obtengan de datos no identificados. (Ribaric, S., Ariyaeeinia, A., y Pavesic, N., 2016).

Para comprender qué tipo de información personal y sensible se puede encontrar en medios audiovisuales, más allá de los marcadores visuales para la identificación física de la persona que se podrían encontrar en la imagen bidimensional (fotografía, plano de video), es importante tener en cuenta las posibilidades que ofrece este tipo de información no estructurada:

10.1 Análisis de audio

El análisis de audio es el proceso de comprimir los datos y empaquetarlos en un solo formato llamado audio. Audio Analytics se refiere a la extracción de significado e información de señales de audio para su análisis.

Hay dos formas de representar el audio: 1) Representación del sonido 2) Archivos de sonido sin formato. El formato de archivo de audio es un formato para datos de audio digital en el sistema. Hay tres formatos de audio principales: formato de audio sin comprimir, formato de audio comprimido sin pérdida, formato de audio comprimido Lossy.

Figura 4. Ejemplo: distorsión de la voz en proceso de anonimización.

10.1.1. Área de aplicación del análisis de audio

El audio es el archivo que se utilizó para transferir los datos de un lugar a otro. El análisis de audio se usa para verificar si los datos del audio están disponibles en el formato adecuado o en un formato similar para el envío del remitente. Las aplicaciones del análisis de audio son muchas:

a. **Aplicación de vigilancia:** la aplicación de vigilancia se basa en un enfoque para la elección sistemática de tipos de audio para la detección de delitos cometidos en la sociedad y en ocasiones, es la única forma de detectar un tipo de actividad sospechosa. La aplicación también se utiliza para enviar información importante sobre vigilancia en alguna situación de crisis.

b. **Detección de amenazas:** el mecanismo de audio se utiliza para identificar el hilo que tiene lugar entre el emisor y el receptor.

c. **Sistema de Tele-monitoreo:** recientes tecnologías en las cámaras de vigilancia incluyen la posibilidad de grabar el audio también. El análisis de audio puede proporcionar una detección efectiva de gritos, vidrios rotos, sonido de arma de fuego, explosiones, llamadas de ayuda, etc. La combinación de análisis de audio y análisis de video en sistemas de monitoreo individuales resulta un medio efectivo de detección de amenazas.

d. **Sistema de red móvil:** el sistema de red móvil se usa para hablar o transferir información de un lugar a otro. A veces, debido a algún problema de red, el sonido de audio no funciona correctamente en ese momento. El análisis de audio se utiliza para buscar la información que no se envía correctamente debido a algunos problemas.

10.2 Análisis de video

El video es un problema importante al considerar Big Data. Los videos y las imágenes representan el 80% de los datos no estructurados. Hoy en día, las cámaras CCTV son la única forma de información y vigilancia digital. Toda esta información se almacena y procesa para su uso posterior, pero el video contiene mucha información y generalmente es de gran tamaño. Por ejemplo, YouTube tiene innumerables videos subidos cada minuto que contienen información masiva. No todos los videos son importantes y no se ven en gran medida, contribuyendo a la generación de "contenidos chatarra", que ocasionan muchos problemas en el análisis de big data. Además de los videos, las cámaras de vigilancia producen mucha información en segundos. Incluso una pequeña cámara digital que captura una imagen almacena millones de píxeles de información en milisegundos.

Dimensiones del Análisis de Datos de Video - Volumen: el tamaño del video al ser mayor, toma tanto de la red como del servidor, en tiempo de procesamiento. Las conexiones de bajo ancho de banda crean mayor tráfico en la red ya que estos videos circulan lentamente. Cuando se almacena en almacenamiento secundario, se requiere una gran cantidad de espacio y se necesita más tiempo para recuperarlo y procesarlo. -Variedad: videos producidos en varios formatos como videos HD, copias de Blu-Ray, etc. -Velocidad: es la velocidad de transmisión de los datos. Hoy en día, las cámaras digitales procesan y capturan videos a muy alta calidad y velocidad. La edición de video hace que crezca en tamaño ya que contienen información adicional sobre los mismos.

Ejemplo: anonimización en video. Fuente: freepik (2453518).

10.2.1 Área de aplicación del análisis de video

- **Útil en casos de accidentes:** con el uso de cámaras de CCTV podemos identificar lo que sucedió en el momento de un accidente. También se utiliza por razones de seguridad como en estacionamiento de vehículos, etc.
- **Útil en las escuelas, policía de tráfico, negocios, seguridad, etc.**
- **Video análisis para la investigación** *(Video Search)***:** se implementan algoritmos de análisis de video para analizar y representa una tarea desafiante y que consume mucho tiempo para el operador humano, especialmente cuando hay una gran cantidad de datos disponibles que podemos buscar en particular cuando lo requerimos.
- **Video análisis para Inteligencia Empresarial** *(Business Intelligence)***:** se utiliza para extraer datos estadísticos y operacionales. En lugar de tener un operador que revise todo el video y haga un recuento de todas las personas o automóviles que se mueven en determinada área, o verifique qué rutas de tráfico se toman con más frecuencia, los análisis de video pueden hacerlo automáticamente.
- **Análisis de objetivos y escenarios:** el análisis de video para inteligencia empresarial involucra análisis de objetivos y escenarios. Target Analytics proporcionan información detallada sobre el movimiento del objetivo, los patrones, la apariencia y otras características que pueden utilizarse para la identificación del objetivo.
- **Direction Analytics:** es la capacidad de distinguir el comportamiento [de las personas] asignando valores específicos (de menor a mayor) a áreas dentro del campo de visión de una cámara.
- **Eliminar la ecuación humana a través de la automatización:** elimina el tedio involucrado en dar uno o más ojos a un monitor por un período prolongado de tiempo. La automatización del análisis de video permite la inserción del juicio humano en el momento más crítico del proceso de vigilancia.

10.3. Elementos a tener en cuenta para la anonimización de datos en formato audiovisual

10.3.1 Formatos de imagen o archivo de audio

Los formatos de imagen son más difíciles de "limpiar" y requieren diferentes técnicas de desidentificación según la información solicitada. Los objetos de archivo de imagen son más complejos y ricos en varios identificadores más allá del nombre o número de teléfono habitual. Además, contienen diferentes formatos que vienen comprimidos o usan gráficos vectoriales alternativos, y hacen que el acceso a archivos de imagen o audio sea aún más turbio. Algunos formatos de imagen también producen metadatos, que, además de PII visual, producen otros indicadores PII, como geolocalización o información en archivos Exif. Crear un proceso automatizado para identificar PII en estos formatos no estructurados es difícil porque requeriría un amplio conocimiento sobre dichos formatos de imagen y los datos binarios producidos. Sin embargo, varios formatos multimedia, como los archivos de imágenes, pueden ser accesibles para ciertas solicitudes en las que el algoritmo de los solicitantes busca metadatos de archivos de imágenes y produce resultados basados en texto para facilitar el borrado de PII.

10.3.2 Otros formatos multimedia

Si bien los archivos de imagen se consideran multimedia, separamos categóricamente las imágenes del video porque existen complejidades adicionales que requerirían una desidentificación multimodal. Agregar otra modalidad aumenta la cantidad de identificadores personales. Además, si la modalidad no se entiende bien, más indicadores personales serían vulnerables a una desidentificación insuficiente y, en última instancia, a una nueva identificación mediante la posible vinculación entre modos múltiples. Por lo tanto, los archivos multimedia presentan un riesgo significativamente mayor para la privacidad de los sujetos, y la accesibilidad debe restringirse.

Ejemplo: análisis de videos. Fuente: freepik (4792593).

Los marcadores PII para los contenidos audiovisuales multimedia, generalmente son contextuales y proveen información de actividades y comportamientos que se pueden clasificar en rangos:

> En función de los tipos de identificadores personales anteriores, [es decir, complementarios] se pueden clasificar en I) identificadores no biométricos, incluidos el contexto del texto, el contexto del discurso, la matrícula [de vehículos e inmuebles], el contexto sociopolítico y ambiental específico, el estilo de vestir y el peinado; II) Los identificadores biométricos son las características personales distintivas, mensurables, generalmente únicas y permanentes que se utilizan para identificar a las personas. A continuación, se clasifican generalmente como fisiológicos (cara, iris, oído, huella digital) frente a conductuales (voz, marcha, gesto, movimiento de labios, estilo de mecanografía), III) Los identificadores biométricos suaves proporcionan características físicas, de comportamiento o adheridas, vagas, que no son necesariamente permanentes o distintivas (altura, peso, color de ojos, silueta, edad, sexo, raza, lunares, tatuajes, marcas de nacimiento, cicatrices) … En la mayoría de los casos, los identificadores biométricos blandos por sí solos no pueden proporcionar una identificación personal confiable, pero pueden usarse para mejorar el rendimiento del reconocimiento, o para clasificar a las personas en categorías particulares, lo que también es intrusivo para la privacidad. (Ribaric, 2016)

10.4 Recomendaciones para la desidentificación y anonimización de información personal en datos audiovisuales y multimedia

Debido a las características particulares de la producción audiovisual y al amplio rango de marcadores de información personal que se puede encontrar en este material, este proceso no puede considerarse como unidireccional, ya que implica el entendimiento del proceso comunicativo que caracteriza este tipo de lenguaje. De la misma forma en que lo audiovisual para su creación necesita incluir diferentes tipos de insumos (lo sonoro, lo visual, lo espacial, temporal, comportamental, etc.), el proceso de desidentificación debe reconocer estos elementos por lo que tendría que ser un proceso igualmente interdisciplinario.

Hasta el momento no se han identificado herramientas informáticas que por sí mismas (automáticamente) realicen un proceso de desidentificación y/o anonimización total de los datos audiovisuales y multimedia. Aunque existen muchos avances con herramientas como la inteligencia artificial, el reconocimiento facial[26] y el análisis de sonidos con algoritmos para restituir los marcadores de identificación (proceso que puede aplicarse a la inversa); hasta el momento lo posible es utilizar estrategias de desidentificación multimodal en los que se pueden disgregar esos identificadores de lo visual y lo sonoro, lo que en realidad significa acciones adicionales en el proceso de posproducción.

Ejemplo: análisis de datos. Fuente: freepik (16312843).

De esta manera, la recomendación principal para este proceso es tener en cuenta las acciones necesarias para la protección de datos personales desde la pre producción del contenido audiovisual, previendo qué tipo de información personal y sensible puede resultar del proceso de producción o recopilación de material, la necesidad de material de apoyo como autorizaciones de uso, sesiones de derechos, locaciones y/o dramatizaciones incluso; así como una construcción de contenidos informada de las restricciones y deberes de autores y productores, similar a la que se utiliza para documentales y noticieros, por ejemplo.

En el caso del material ya grabado, crudo o editado, los procesos de posproducción es en su mayoría lo recomendado para desvincular los marcadores de identificación de los sujetos en el material audiovisual. Aunque pueda parecerlo, las distorsiones de la imagen, la luz y el sonido no son suficientes debido a la gran cantidad de información que puede captar una grabación audiovisual y/o sonora. Por esta razón se emplean con frecuencia locaciones “neutras” y/o utilería y escenificación poco reconocibles en los procesos de producción. Para la posproducción entonces, tener en cuenta el sonido circundante, incidental y ambiental en el caso del audio resulta fundamental; así como el espacio, las características de la utilería, espacios, indicaciones climáticas e incluso ambientales en los registros fotográfico y de video, son aspectos que se deben tratar con las acciones de edición necesarias, además de borrar o reemplazar cualquier dato de información personal que pueda contener el material.

Aunque actualmente este campo está en pleno desarrollo, “A pesar de los enormes esfuerzos de diversos grupos académicos de investigación, instituciones y empresas, la investigación en el campo de la desidentificación y la desidentificación multimodal en contenido multimedia aún está iniciando. Se ha hecho relativamente poco en el campo de la desidentificación de identificadores no biométricos, excepto en el campo del texto y la desidentificación de matrículas. Para evitar la identificación de "restricción de pares" y la clasificación de individuos en categorías (que pueden tratarse como invasivas a la privacidad), se deben realizar esfuerzos adicionales en el campo del estilo de vestirse y la identificación del peinado (inicialmente se han realizado esfuerzos solo para ocultar los peinados y el color del cabello). El problema de ocultar o eliminar selectivamente la información sensible al contexto u objetos del entorno que pueden usarse para revelar la identidad de una persona aún está abierta. Esto podría resolverse en el futuro cercano utilizando un enfoque basado en el conocimiento para modelar un entorno y situación específicos para detectar ROI adicionales y oscurecerlos”.

10.4.1 Normas a tener en cuenta

Directiva 95/46/CE del Parlamento Europeo y del Consejo de 24 de octubre de 1995 relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (RGPDUE)”.

REFERENCIAS

- Agencia Española para la protección de datos. (2016). Orientaciones y garantías en los procedimientos de anonimización de datos personales. Guía de lineamientos. https://www.aepd.es/sites/default/files/2019-12/guia-orientaciones-procedimientos-anonimizacion.pdf
- Arias, P., Soladie, C., Bouafif, O., Roebel, Axel., y Seguier, R. (2020). Realistic transformation of facial and vocal smiles in real-time audiovisual streams. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8307228
- Clunie, D. (2021). Un software de libre acceso y de uso relativamente sencillo es recomendado por la Universidad Johns Hopkins para la "limpieza" de PII en registros de imagen en historias clínicas de pacientes. http://www.dclunie.com/
- Congreso de la República (Julio 14, 2000). Ley 594 de 2000. Ley General de Archivos. https://normativa.archivogeneral.gov.co/ley-594-de-2000/
- Congreso de la República. (Marzo 6, 2014). Ley 1712 de 2014. Transparencia y del Derecho de Acceso a la Información Pública Nacional. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=5688
- Congreso de la República. (Octubre 17, 2012). Ley estatutaria 1581 de 2012. Disposiciones generales para la protección de datos personales. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981
- Consejo Internacional de Archivos. (2014). Guía Técnica para la Gestión de Archivos de Uso Restringido. https://www.ica.org/es/guia-tecnica-para-la-gestion-de-archivos-de-uso-restringido
- Consejo Nacional de Política Económica y Social (CONPES). (2018). Política Nacional de Explotación de Datos (Big Data). Departamento Nacional de Planeación, Ministerio de Tecnologías de la Información y las Comunicaciones, Superintendencia de Industria y Comercio.
- Corte Constitucional de Colombia. (Marzo 17, 2003). Sentencia 227. https://www.corteconstitucional.gov.co/relatoria/2003/t-227-03.htm
- Departamento Administrativo Nacional de Estadística. (2014). Lineamientos para la anonimización de microdatos. DANE. https://www.dane.gov.co/files/sen/registros-administrativos/guia-metadatos.pdf
- Departamento Administrativo Nacional de Estadística. (2018). Colombia – Encuesta Anual de Comercio - EAC - 2016. DANE. http://microdatos.dane.gov.co/index.php/catalog/520
- Grupo de trabajo sobre protección de datos. (2014). Dictamen 05 / 2014 sobre técnicas de anonimización. Comisión Europea. https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf
- International Organization for Standardization. (2011). Estandar ISO / IEC 29100:2011. Unión Europea: Organización Internacional de Normalización.
- Jaffré, G., y Pinquier, J. (2021). Audio/Video Fusion: a Preprocessing Step for Multimodal Person Identification. Universidad Paul Sabatier. http://mmua.cs.ucsb.edu/MMUA2006/Papers/116.pdf
- Luque, J., Morros, R., Garde, I., Anguita, J., Farrús, M., Macho, D., Marqués, F., Martínez, C., Vilaplana, V., y Hernando J. (2007) Audio, video and multimodal person identification in a smart room. https://repositori.upf.edu/handle/10230/32740?locale-attribute=es
- Ministerio de Salud y Protección Social. (s.f.). Lineamientos para la anonimización de datos del sistema nacional de estudios y encuestas poblacionales para la salud. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/lineamientos-anonimizacion-sistema-encuestas.pdf
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2016). Guía dedatos abiertos en Colombia. https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf
- Octave. (2011). Iniciativa europea para la verificación automática de altavoz. https://www.octave-project.eu/
- Personal Data Protection Commission, Singapore. (2018). Guide to Basic Data Anonymization Techniques. https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/Guide-to-Anonymisation_v1-(250118).pdf
- Presidencia de la República de Colombia. (Junio 27, 2013). Decreto 1377. Reglamentación parcial de la Ley 1581 de 2012.
- Real Academia Española. (2017). Diccionario de la Lengua Española. http://dle.rae.es
- Ribaric, S., Ariyaeeinia, A., y Pavesic, N. (2016). De-identification for privacy protection in multimedia content: A survey. https://www.researchgate.net/publication/303775509_De-identification_for_privacy_protection_in_multimedia_content_A_survey
- Sánchez, D., Soria-Comas, J., & Domingo-Ferrer, J. (2016). Database anonymization: privacy models, data utlity, and microaggregation-based inter-model connections. Information security, privacy and Trust, 4.
- Secretaría de Transparencia de la Presidencia de la República. (2016). Guía de instrumentos de información pública. https://dapre.presidencia.gov.co/AtencionCiudadana/DocumentosCiudadania/guia-instrumentos-gestion-informacion-publica-SecTransparencia-DAPRE.pdf
- Shah, D., Han, K., y Narayanan, S. (2010). Robust multimodal person recognition using lowcomplexity audio-visual feature fusion approaches. https://sail.usc.edu/publications/files/ShahHanNarayanan-IJSC-2010.pdf

ANEXOS

1. TABLA DE RESUMEN DE TÉCNICAS DE ANONIMIZACIÓN

NOMBRE DE LA TÉCNICA	CUÁNDO USAR	TIPO DE ATRIBUTO
Supresión de atributos	El atributo no es necesario en el conjunto de datos anonimizados	Todos
Supresión de registros	Presencia de registros atípicos	N.A. (se aplica a todo el registro, de ahí todos los atributos afectados)
Enmascaramiento de caracteres	Enmascarar algunos caracteres en un atributo proporciona suficiente anonimato	Identificador directo
Seudonimización	Los registros aún deben distinguirse entre sí en el conjunto de datos anónimo, pero no se puede retener ninguna parte del valor del atributo original	Identificador directo
Generalización	Los atributos se pueden modificar para ser menos precisos pero aún así ser útiles	Todos
Intercambio	No es necesario analizar las relaciones entre los atributos en el nivel de registro	Todos
Perturbación de datos	Leve modificación a los atributos son aceptables	Identificador indirecto
Datos sintéticos	Se requieren grandes cantidades de datos inventados de naturaleza similar a los datos originales, p.ej. para pruebas de sistema	Todos
Agregación de datos	No se requieren registros individuales y los datos agregados son suficientes	Identificador indirecto

2. NORMOGRAMA

A continuación, se presentan las normas nacionales relacionadas con los temas de esta guía.

GUÍA DE ANONIMIZACIÓN NORMATIVIDAD PERTINENTE			
TIPO DE NORMA	NÚMERO	AÑO	DESCRIPCIÓN
Constitución Política de Colombia		1991	Artículo 15. Todas las personas tienen derecho a su intimidad personal y familiar y a su buen nombre, y el Estado debe respetarlos y hacerlos respetar. De igual modo, tienen derecho a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bancos de datos y en archivos de entidades públicas y privadas.
Ley	1266	2008	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.
Ley	1581	2012	Por la cual se dictan disposiciones generales para la protección de datos personales.
Ley	1712	2014	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.
Ley	1755	2015	Por medio de la cual se regula el Derecho Fundamental de Petición y se sustituye un título del Código de Procedimiento Administrativo y de lo Contencioso Administrativo.
Ley	1273	2009	Por medio de la cual se modifica el Código Penal, se crea un nuevo bien jurídico tutelado - denominado "de la protección de la información y de los datos" - y se preservan integralmente los sistemas que utilicen las tecnologías de la información y las comunicaciones, entre otras disposiciones.
Lety	527	1999	Por medio de la cual se define y reglamenta el acceso y uso de los mensajes de datos, del comercio electrónico y de las firmas digitales, y se establecen las entidades de certificación y se dictan otras disposiciones.
Ley	594	2000	Por medio de la cual se dicta la Ley General de Archivos y se dictan otras disposiciones.
Decreto	235	2010	Por el cual se regula el intercambio de información entre entidades para el cumplimiento de funciones públicas.
Decreto	2280	2010	Por el cual se modifica el artículo 3° del Decreto 235 de 2010.
Decreto	1377	2013	Reglamenta parcialmente la Ley 1581 de 2012.

GUÍA DE ANONIMIZACIÓN NORMATIVIDAD PERTINENTE			
TIPO DE NORMA	**NÚMERO**	**AÑO**	**DESCRIPCIÓN**
Decreto	103	2015	Disposiciones generales en materia de transparencia y del derecho de acceso a la información pública nacional.
Decreto	415	2016	Por el cual se adiciona el Decreto Único Reglamentario del sector de la Función Pública, Decreto Numero 1083 de 2015, en lo relacionado con la definición de los lineamientos para el fortalecimiento institucional en materia de tecnologías de la información y las comunicaciones.
Decreto	3851	2006	Por el cual se organiza un sistema de aseguramiento de la calidad, almacenamiento y consulta de la información básica colombiana y se dictan otras disposiciones.
Decreto	32	2013	Por el cual se crea la Comisión Nacional Digital y de Información Estatal.
Decreto	4800	2011	Por el cual se reglamenta la Ley 1448 de 2011 y se dictan otras disposiciones.
Decreto	2758	2012	Por el cual se reglamenta el Sistema Nacional de Archivos, se establece la Red Nacional de Archivos, se deroga el Decreto número 4124 de 2004 y se dictan otras disposiciones relativas a la administración de los archivos del Estado.
Decreto	2609	2012	Por el 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".
Decreto	1413	2017	Por el cual se adiciona el Título 17 a la Parte 2 del Libro 2 del Decreto Único Reglamentario del sector de Tecnologías de la Información y las Comunicaciones, Decreto número 1078 de 2015, para reglamentarse parcialmente el Capítulo IV del Título III de la Ley 1437 de 2011 y el artículo 45 de la Ley 1753 de 2015, estableciendo lineamientos generales en el uso y operación de los servicios ciudadanos digitales.
Decreto	1828	2017	Por el cual se crea el Sistema Integrado de Información para el Posconflicto - SIIPO.
Directiva Presidencial	2	2000	Plan de Acción de la Estrategia de Gobierno en Línea.
Directiva Presidencial	4	2012	Eficiencia Administrativa y lineamientos de la política cero papel en la administración pública.
Resolución	3564	2015	Por la cual se reglamentan los artículos 2.1.1.2.1.1, 2.1.1.2.1.11, 2.1.1.2.2.2, y el parágrafo 2 del artículo 2.1.1.3.1.1 del Decreto N° 1081 de 2015.

GUÍA DE ANONIMIZACIÓN NORMATIVIDAD PERTINENTE			
TIPO DE NORMA	**NÚMERO**	**AÑO**	**DESCRIPCIÓN**
Circular	58	2009	Cumplimiento Decreto 1151 de 2008.
Sentencia	C 1011	2008	Revisión de constitucionalidad del Proyecto de Ley Estatutaria No. 27/06 Senado - 221/07 Cámara (Acum. 05/06 Senado).
Sentencia	C 274	2013	Revisión constitucional del Proyecto de Ley Estatutaria número 228 de 2012 Cámara, 156 de 2011 Senado, "por medio de la cual se crea la ley de transparencia y del derecho de acceso a la información pública nacional".
Sentencia	C 951	2014	Revisión de constitucionalidad del Proyecto de Ley número 65 de 2012 Senado y número 227 de 2013 Cámara "Por medio del cual se regula el derecho fundamental de petición y se sustituye un título del Código de Procedimiento Administrativo y de lo Contencioso Administrativo".
Sentencia	C 729	2002	Acción de tutela instaurada por Carlos Antonio Ruiz Gómez contra el Departamento Administrativo de Catastro (Alcaldía Mayor de Bogotá) y la Superintendencia Nacional de Salud.
Acuerdo	5	2013	Por el cual se establecen los criterios básicos paro la clasificación, ordenación y descripción de los archivos en las entidades públicas y privadas que cumplen funciones públicas y se dictan otras disposiciones.

GUÍA PARA LA ANONIMIZACIÓN DE DATOS E INFORMACIÓN NO ESTRUCTURADA:

ESTÁNDARES Y LINEAMIENTOS TÉCNICOS

Centro Nacional de Memoria Histórica