



UAECD

Buenas Prácticas para la Publicación de Datos.

—

Buenas prácticas para la publicación de datos

Fecha de creación: septiembre de 2018

Página web: www.ideca.gov.co

Correo electrónico: ideca@catastro.gov.co

Licencia: Attribution 4.0 International (CC BY 4.0)

Autores: Unidad Administrativa Especial de Catastro Distrital -
Gerencia IDECA



Control De Versiones

| Fecha | Autor/ Modificado por | Versión | Cambio efectuado |
|------------|-------------------------------|---------|---|
| 2018-10-22 | Martha Elisabeth Melo Beltrán | 1.0 | Primera versión del documento. No hay cambios para registrar. |
| 2019-07-05 | Astrid Yadira Duitama Guio | 1.1 | Ajustes Generales, Cambio de formato. |



Contenido

| | |
|--|----|
| Introducción..... | 5 |
| Objetivo..... | 7 |
| Audiencia..... | 7 |
| Alcance..... | 7 |
| 1. Proporcionar metadatos asociados a los datos publicados..... | 9 |
| 2. Proporcionar licencia en donde se especifiquen las limitaciones y restricciones de acceso y uso de los datos..... | 13 |
| 3. Utilizar vocabularios estandarizados y reutilizables en la estructuración y publicación de los datos..... | 16 |
| 4. Anonimizar los datos de carácter personal, información estadística, información indirectamente identificable e información sensible o confidencial..... | 20 |
| 5. Publicar los datos en formatos estándar, reutilizables y legibles por aplicaciones informáticas..... | 23 |
| 6. Proporcionar un diccionario de datos que facilite el uso e interpretación de los datos publicados..... | 26 |
| 7. Acompañar la publicación de los datos con enlaces a otros recursos..... | 28 |
| 8. Depurar sus datos antes de publicarlos para garantizar su calidad y aumentar su confianza..... | 30 |
| 9. Incorporar el componente geográfico en los datos publicados para optimizar su comprensión, uso, re-uso y visualización..... | 34 |
| 10. Asignar un identificador único que permita aumentar la consulta, uso e interoperabilidad del dato publicado..... | 37 |
| Obtenga el más Alto Nivel en Publicación de Datos..... | 39 |
| Referencias..... | 40 |



Introducción

La Infraestructura de Datos Espaciales del Distrito Capital - IDECA es el instrumento que permite a las autoridades del Distrito identificar de manera constante la condición, movilidad, dinámica, expansión, proceso, proyectos y necesidades de la ciudad de Bogotá a través de la obtención, uso y disposición de información georreferenciada para atender a la ciudadanía y contribuir en la toma de decisiones en beneficio de la ciudad y de sus habitantes.

La producción de la información está a cargo de las entidades aportantes, en desarrollo de sus actividades misionales y el uso de los datos producidos dependerá del propósito del dato, de la finalidad específica del usuario y de cómo y cuando estén disponibles los datos.

IDECA gestiona la disposición de los datos, lo cual implica procurar que los datos publicados por las entidades pertenecientes a la Administración Distrital cuenten con altos estándares de calidad, interoperabilidad, documentación y promuevan y faciliten el acceso, comprensión, uso y re-uso de estos por cualquier usuario.

Para encontrar y acceder los datos disponibles en el entorno Web, los usuarios no especializados tienen algunas dificultades, ya que encuentran contenidos duplicados, desactualizados e inconsistentes; lo que ha generado para los publicadores el reto de garantizar el descubrimiento, acceso, calidad e interoperabilidad de sus datos.

Es importante mencionar que, si se publican datos útiles, novedosos, actualizados y de calidad, no solo aumenta el número de usuarios que los reutilicen y distribuyan, sino que aumenta la credibilidad y confianza en los datos. Adicionalmente se disminuyen los costos y tiempos en la generación de los datos y se reducen los usos indebidos de los mismos.

Esta guía pretende que las entidades o responsables de publicar y disponer los datos lo hagan con altos estándares de calidad, debidamente documentados, en formatos interoperables y que satisfagan necesidades reales de los usuarios para que estos últimos puedan encontrarlos, accederlos, comprenderlos y reutilizarlos.



Este documento reúne las mejores prácticas de publicación de datos en la Web, cubriendo aspectos como el formato, el acceso, la calidad, los metadatos, el licenciamiento y el origen de los datos.

Se invita al lector a navegar fácilmente en el documento, ya que este presenta sin un orden determinado las mejores prácticas de publicación de datos explicadas en:



¿Qué logro?: Los beneficios que se obtienen al aplicar la buena práctica;



¡Recomendaciones!: Las herramientas básicas para conseguir el objetivo de la buena práctica;



Ejemplos: Utilizando conjuntos de datos reales de las entidades oficiales del Distrito Capital se visualiza la aplicación de la buena práctica.



Objetivo

Promover en las entidades de la Administración Distrital y sus responsables, la apropiación y adopción de las mejores prácticas de publicación de datos con el fin de aumentar el valor en la información producida dentro de su quehacer institucional.

Audiencia

Este documento está dirigido a las entidades y personas responsables de la Administración Distrital que publican o desean publicar datos oficiales de Bogotá a través de cualquier plataforma de datos.

Alcance

Las mejores prácticas expuestas en este documento están encaminadas a que los datos publicados a través de plataformas oficiales cumplan con unos requisitos mínimos de calidad, documentación interoperabilidad, y licenciamiento, dentro del proceso de publicación.

Buscamos con esta guía de buenas prácticas garantizar la apropiación de cuatro principios:

- ✓ Que los datos sean descubiertos, es decir que cualquier usuario en el entorno Web los pueda encontrar fácilmente;
- ✓ Que los datos sean accesibles, quiere decir que no presenten restricciones de acceso ni que se encuentren cifrados;
- ✓ Que los datos sean interoperables, en otras palabras, que sean publicados en formatos estandarizados que permitan ser procesados por cualquier software;
- ✓ Que los datos sean reutilizables, es decir que estos sean utilizados en la mayor cantidad de productos, servicios y mercados.





Debe tenerse en cuenta que IDECA expone en este documento las mejores prácticas de publicación de datos, identificadas con base a la experiencia adquirida en la disposición datos de los últimos años, sin embargo, existe una amplia gama de otras buenas prácticas de publicación que no han sido incluidas en el presente texto, y en consecuencia se invita al lector a apropiarse de las prácticas expuestas y a dejar abierta la posibilidad de complementar la información presentada con otras fuentes.



Buenas Prácticas para la Publicación de Datos

1. Proporcionar metadatos asociados a los datos publicados, que permitan conocer las características de los datos y sean interpretables tanto por usuarios como por aplicaciones informáticas.

Un metadato, entendido como la información acerca de un recurso¹, es un dato con muchos elementos que le permiten al usuario conocer desde la visión del productor, las características propias de los datos publicados, tales como las fechas de creación y actualización, el canal de acceso, la calidad del dato, las condiciones de uso, quién lo produjo, quién lo distribuye, etc. Esto quiere decir que el metadato explica los datos, conteniendo toda la información relevante del dato.

En la vida cotidiana, la cédula de ciudadanía es el documento que contiene la información mínima de las personas (nombre, sexo, fecha de nacimiento, G.S. RH, estatura, lugar de nacimiento); esta información permite que el ciudadano se identifique en cualquier lugar del territorio nacional ante otras personas o autoridades que requieran conocerlo. Con los datos el tratamiento es similar. El dato correspondería a la persona y el metadato corresponde a la cédula de ciudadanía. Así como en Colombia todo ciudadano debe portar su documento de identificación, en el entorno Web todo dato debe contar con metadatos completos, en otras palabras, el metadato es parte integral del dato y debe existir entre ambos un vínculo inseparable.

Los datos siempre deberán publicarse con su respectivo metadato, este último deberá contar con información completa y estandarizada que garantice que los usuarios encuentren, conozcan y utilicen los datos que se disponen.

¹ ISO 19115-1:2014, 4.10





Recuerde que la publicación del metadato es igual de importante que la publicación del dato.



¿Qué logro?

Al publicar datos en conjunto con su metadato obtengo:

- ✓ Que los datos sean encontrados y utilizados por más usuarios.
- ✓ Que los usuarios logren una correcta interpretación de la naturaleza, calidad y distribución de los datos
- ✓ Que los datos sean entendidos, evaluados y retroalimentados por más usuarios.
- ✓ Que los datos puedan ser compartidos e intercambiados entre sistemas.
- ✓ Que los motores de búsqueda encuentren y refieran a los datos que se han dispuesto.
- ✓ Que las aplicaciones informáticas procesen mejor y en mayor cantidad mis datos.



¡Recomendaciones!

- ✓ Disponga metadatos suficientes que permitan a los usuarios tener la información correcta para decidir si el dato publicado satisface o no sus necesidades.
- ✓ Incluya en sus metadatos la ubicación del dato, a través de la descripción del lugar (localidad, parque, río, etc.) o a través de localización por coordenadas geográficas x e y.
- ✓ Publique sus metadatos como una página Web HTML (HyperText Markup Language) o simplemente como un archivo de texto separado.
- ✓ Disponga sus metadatos en el formato básico de intercambio XML (Extensible Markup Language) o en formatos estándar más sofisticados como JSON (JavaScript Object Notation).
- ✓ Al igual que los datos, puede publicar los metadatos en múltiples formatos desde la misma URL.
- ✓ Tenga en cuenta que IDECA a partir de estándares internacionales ha definido los elementos mínimos que debe contener el metadato y que estos al ser utilizados en la publicación, facilitan el intercambio de los datos.



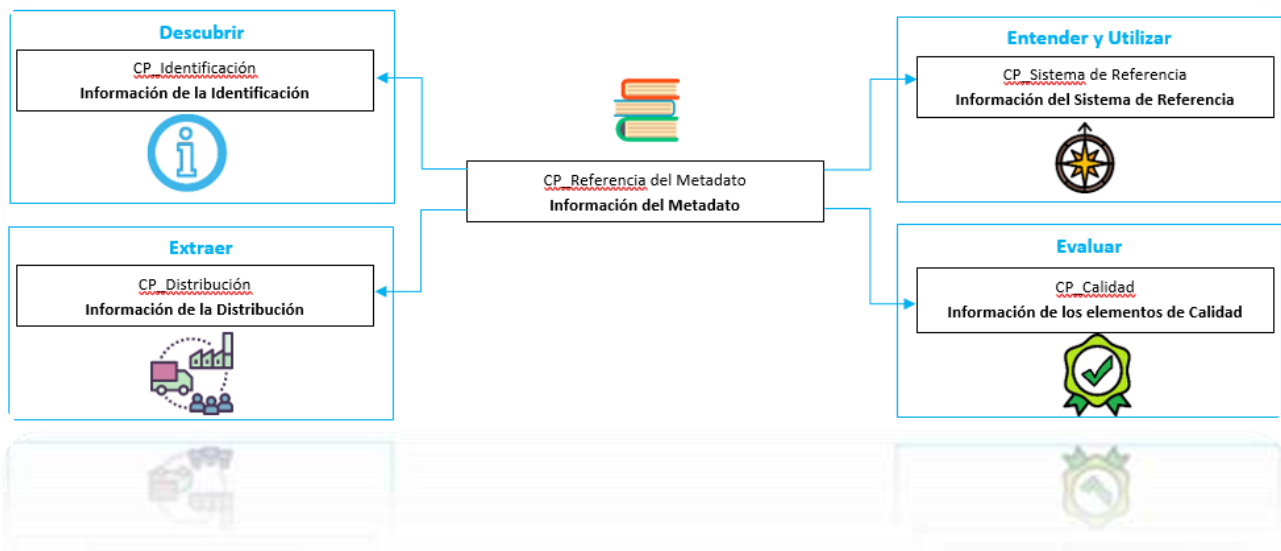



Figura 1: Secciones del Perfil de Metadatos IDECA

 Para más información consulte la [“Cartilla para la creación de metadatos geográficos en Bogotá Distrito Capital”](#).

 Visite la sección [“Construye el metadato”](#) y aprenda a crear fácilmente su metadato.

 **Ejemplo 1:**

En la siguiente URL se observa un metadato dispuesto para que los usuarios puedan consultar fácilmente la información asociada al dato publicado (Áreas Compatibles con Minería de la Secretaría Distrital de Ambiente).

<http://datosabiertos.bogota.gov.co/dataset/areas-compatibles-con-mineria>

El metadato es fácilmente consultable en la sección “Información Adicional”



| Información Adicional | |
|------------------------------------|--|
| Formato de presentación | Mapa Digital |
| Frecuencia de Actualización | Irregular |
| Palabras clave | Estrategia, Minería, Sostenibilidad Ambiental |
| Enlace a otros recursos | http://www.ideca.gov.co http://www.ideca.gov.co/geolocalizacion |
| Tipo de representación espacial | Vector |
| Resolución Espacial | 100 |
| Tipo de Sistema de Referencia | EPSG:4886 MATDRA-SRSDA0 geographic ID - 4886 |
| Categoría Temática | Ambiente |
| Ubana | Suárez |
| Alcance de la calidad de los datos | Conjunto de Datos |
| Reporte de la calidad | Descripción del procedimiento de creación de la capa: Fuentes: Los datos geográficos ambientales corresponden a la información integrada en esta capa de información, correspondiente a los niveles de los Áreas Corregidas con Minería, obtenidos mediante la Resolución MADT 2301 de 2010 - Artículo 7º. Puede ser aplicable diferentes entre a Nivel Nacional, Regional y Institucional, si |



Ejemplo 2:

En la siguiente URL está disponible el metadato de un nivel de información dispuesto por la Unidad Administrativa Especial de Catastro Distrital - UAECD - en formato XML (legible por aplicaciones informáticas).

<http://datosabiertos.bogota.gov.co/metadato/dataset?id=397ccbdb8-e2c5-4700-b90e-b68d101ab0c5>

```

<gmd:fileMetadata xsi:nil="http://www.isotc211.org/2005/gmd" xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gv="http://www.isotc211.org/2005/gv" xmlns:gd="http://www.ideca.gov.co/1.0">
  <gmd:fileIdentifier>
    <gmd:CharacterString>[8f7ccbb8-e2c5-4700-b90e-b68d101ab0c5]</gmd:CharacterString>
  </gmd:fileIdentifier>
  <gmd:language>
    <gmd:CharacterString>es</gmd:CharacterString>
  </gmd:language>
  <gmd:characterSet>
    <gmd:MD_CharacterSetCode codeSpace="ISO10646/UCS2" codeList="http://www.isotc211.org/2005/resources/CodeLists/gmdCharacterSetCodeCodeList.htm" codeListUse="utf8"/>
  </gmd:characterSet>
  <gmd:fileIdentifier>
    <gmd:MD_ScopeCode codeSpace="ISO15926/19118" codeList="http://www.isotc211.org/2005/resources/CodeLists/gmdCodeLists.nilMD_ScopeCodeCodeList.htm" dataURI="dataset"/>
  </gmd:fileIdentifier>
  <gmd:contact>
    <gmd:CI_ResponsibleParty>
      <gmd:organisationName>
        <gmd:CharacterString>
          Unidad Administrativa Especial de Catastro Distrital (UAECD)
        </gmd:CharacterString>
        <gmd:organisationName>
        </gmd:organisationName>
      </gmd:organisationName>
      <gmd:contactInfo>
        <gmd:CI_Contact>
          <gmd:address>
            <gmd:CI_Address>
              <gmd:country>
                <gmd:CharacterString>Colombia</gmd:CharacterString>
              </gmd:country>
              <gmd:CI_Address>
              </gmd:CI_Address>
            </gmd:address>
            <gmd:CI_Contact>
              <gmd:contactInfo>
                <gmd:CI_Contact>
                  <gmd:CI_Contact>
                    <gmd:CI_Contact>
                      <gmd:CI_Contact>
                        <gmd:CI_Contact>
                          <gmd:CI_Contact>
                            <gmd:CI_Contact>
                              <gmd:CI_Contact>
                                <gmd:CI_Contact>
                                  <gmd:CI_Contact>
                                </gmd:CI_Contact>
                              </gmd:CI_Contact>
                            </gmd:CI_Contact>
                          </gmd:CI_Contact>
                        </gmd:CI_Contact>
                      </gmd:CI_Contact>
                    </gmd:CI_Contact>
                  </gmd:CI_Contact>
                </gmd:CI_Contact>
              </gmd:CI_Contact>
            </gmd:contactInfo>
          </gmd:CI_Contact>
        </gmd:contactInfo>
      </gmd:ResponsibleParty>
    </gmd:contact>
  </gmd:fileMetadata>
</pre>

```



2. Proporcionar una licencia en donde se especifiquen las limitaciones y/o restricciones de acceso y uso de los datos.

Una licencia, entendida como las condiciones legales bajo las cuales está disponible la obra², es un instrumento mediante el cual el propietario del dato autoriza o no a los usuarios para que accedan a su contenido y también lo faculta para su utilización. Existen diferentes tipos de licencia para la distribución y explotación de obras y cuyo propósito es regular las condiciones de acceso y uso de los datos.

Dependiendo el acceso autorizado a los datos, las licencias pueden clasificarse en:

- ✓ Privativas o restrictivas: Aquellas licencias conocidas como cerradas o no libres, donde la mayoría de las veces solo el autor puede usar la obra, o en caso contrario se definen condiciones específicas que se deben cumplir para hacer uso de la información;
- ✓ Semilibres: Aquellas licencias que requieren autorizaciones, para que los usuarios utilicen, copien, distribuyan o modifiquen su contenido;
- ✓ Libres o abiertas: Aquellas licencias que en términos generales permiten usar, distribuir y compartir los datos, con muy pocas restricciones.

Cuando un dato se publica acompañado de su licencia, el usuario conoce las condiciones de acceso, uso y distribución de los datos que está consumiendo, de esta forma se garantiza que los datos se distribuyan según los intereses de la entidad que los produce y se minimizan los usos indebidos de los datos por parte de los usuarios.

Ahora bien, definir la licencia adecuada para cada dato, dependerá de las condiciones que establezca la entidad productora de este. Revise los ejemplos y defina la licencia más adecuada para sus datos.



Para más información consulte el [Instructivo para la definición de licencias de datos](#).

² Recuperado de <https://opendefinition.org/od/2.1/es/>





¿Qué logro?

- ✓ Que el usuario evalúe si puede o no usar mis datos.
- ✓ Que el usuario conozca las condiciones de acceso, uso, y manejo sobre los datos disponibles.
- ✓ Dar claridad y potenciar el uso de los datos disponibles.
- ✓ Que el usuario de los reconocimientos respectivos a la entidad productora de los datos.
- ✓ Que se genere valor agregado sobre los datos disponibles, en caso de asignar una licencia abierta.



¡Recomendaciones!

- ✓ Identifique la licencia que mejor se acomode a su modelo de negocio. Recuerde que mediante la Ley 1712 de 2014 se reguló el derecho de acceso a la información pública, y se incluyeron los principios de facilitación, gratuidad y divulgación proactiva de la información.
- ✓ Disponga una licencia para su conjunto de datos a través de un enlace, una copia incorporada o dentro del metadato.
- ✓ Prefiera el uso de licencias abiertas, de este modo sus datos serán utilizados en una mayor cantidad de productos, servicios y mercados.



Ejemplo 1:

Mundialmente conocidas las licencias Creative Commons³ tienen como objetivo fortalecer a los creadores para que sean quienes definan los términos en que sus obras pueden ser usadas, qué derechos desean entregar y en qué condiciones lo harán. A continuación, encontrará los enlaces de acceso para conocer las condiciones de cada una de las licencias disponibles:

CCO 1.0

<https://creativecommons.org/publicdomain/zero/1.0/deed.es>

³ <http://creativecommons.org/>



UAECD

| | |
|-----------------|---|
| CC BY 4.0 | https://creativecommons.org/licenses/by/4.0/ |
| CC BY-SA 4.0 | https://creativecommons.org/licenses/by-sa/4.0/ |
| CC BY-ND 4.0 | https://creativecommons.org/licenses/by-nd/4.0/ |
| CC BY-NC 4.0 | https://creativecommons.org/licenses/by-nc/4.0/ |
| CC BY-NC-SA 4.0 | https://creativecommons.org/licenses/by-nc-sa/4.0/ |



Ejemplo 2:

También existen las licencias abiertas para datos Open Data Commons⁴ proporcionadas por la mundialmente conocida red Open Knowledge Foundation Network - OKFN⁵, puede consultar las condiciones ingresando a los siguientes enlaces:

| | |
|------------|---|
| ODbL 1.0 | https://opendatacommons.org/licenses/odbl/summary/ |
| PDDL 1.0 | https://opendatacommons.org/licenses/pddl/summary/ |
| ODC-BY 1.0 | https://opendatacommons.org/licenses/by/summary/ |

⁴ <http://opendatacommons.org/>

⁵ <https://okfn.org/>



3. Utilizar vocabularios estandarizados y reutilizables en la estructuración y publicación de los datos.

Un vocabulario define los conceptos y las relaciones que se utilizan para describir y representar un área de interés⁶. Cuando se publican datos en la Web, los vocabularios permiten integrar datos cuando existen ambigüedades en los términos utilizados en los diferentes conjuntos de datos. En la práctica, un vocabulario es un conjunto de palabras que permiten a dos usuarios comunicarse entre sí. La efectividad de dicha comunicación radica en que el significado de una palabra para el usuario emisor sea el mismo que para el usuario receptor independientemente de si utilizaron el mismo término.

Cuando se habla de interoperabilidad de datos, lo que se espera es que tanto los usuarios como las aplicaciones informáticas entiendan y comprendan la mayor cantidad de información que están consumiendo. Para conseguir esto, las comunidades de datos han identificado vocablos, palabras o términos de uso común, los cuales cuentan con significados apropiados, y que actualmente se encuentran integrados y estructurados en diferentes estándares internacionales.

Se recomienda utilizar vocabularios de este tipo para estandarizar los datos contenidos en la información publicada.



¿Qué logro?

- ✓ Que mis datos sean consultados, utilizados y reutilizados por más usuarios.
- ✓ Que se reduzca la ambigüedad en la interpretación de mis datos.
- ✓ Que mis datos sean interoperables.
- ✓ Que mis datos sean procesados en entornos nacionales e internacionales.
- ✓ Que mis datos puedan ser compartidos e intercambiados entre sistemas.
- ✓ Que mis datos sean más confiables.

⁶ <https://www.w3.org/standards/semanticweb/ontology>





¡Recomendaciones!

- ✓ Consulte normas locales o vocabularios internacionales para la estructuración⁷ de sus datos.
- ✓ Apropie los vocabularios consultados y adapte sus datos a estos en los casos en que sea posible.
- ✓ Incluya vocabularios ya definidos y reconocidos internacionalmente cuando vaya a producir nuevos datos.



Ejemplo 1:

A continuación, se presentan algunos vocabularios internacionales que puede consultar y apropiar para optimizar el intercambio de sus datos:

| | | |
|--|------------|---|
| Data Catalog (DCAT) | Vocabulary | https://www.w3.org/TR/vocab-dcat/#class-catalog |
| FOAF Specification | Vocabulary | http://xmlns.com/foaf/0.1/ |
| ISO/TC 211, Geographic information/Geomatics Terminology | | http://www.isotc211.org/Terminology.htm |
| The RDF Data Cube Vocabulary | | https://www.w3.org/TR/2014/REC-vocab-data-cube-20140116/ |
| ODRL Vocabulary Expression 2.2 | & | https://www.w3.org/TR/2018/REC-odrl-vocab-20180215/ |
| HTTP Vocabulary in RDF 1.0 | | https://www.w3.org/TR/2017/NOTE-HTTP-in-RDF10-20170202/ |
| vCard Ontology for describing People and Organizations | - for and | https://www.w3.org/TR/vcard-rdf/ |

⁷ Estructurar significa ajustar los datos a un esquema fijo.





Ejemplo 2:

En el contexto Nacional, el Ministerio de Tecnologías de la Información y las Comunicaciones ha establecido el Lenguaje Común de Intercambio de como un instrumento de consulta que contiene la definición de los elementos de datos, conceptualizados por las entidades del estado. La consulta la puede realizar en <http://lenguaje.mintic.gov.co/>



Ejemplo 3:

La especificación vCard disponible en <https://www.w3.org/TR/vcard-rdf/> promueve un vocabulario específico para describir personas y organizaciones.

Uno de los vocablos que describe esta especificación es la “dirección de entrega” llamado dentro del vocabulario como **ADR** y definido como “la dirección del objeto representado en partes estructuradas”.



2.4 Delivery Addressing Properties

| RFC Property | Note | Ontology Property | N-Ary Property |
|--------------|---|---|---|
| ADR | The address of the object represented in structured parts | hasAddress (range of class Address) street-address locality region country-name postal-code | hasStreetAddress hasLocality hasRegion hasCountryName hasPostalCode |

Ahora bien, este vocabulario especifica los componentes de la dirección de entrega para el objeto así: el nombre del país, la localidad, el código postal, la región, y la dirección dada por vías (calles y carreras en el caso colombiano). Esto quiere decir, que cualquiera de los componentes descritos puede usarse como dirección de entrega válida ya que son equivalentes entre sí.

Address [back to **TeC** or **Class TeC**](#)

IRI: <http://www.w3.org/2006/vcard/ns#Address>

is defined by
<http://www.w3.org/2006/vcard/ns>

To specify the components of the delivery address for the object

is equivalent to
 $((\text{country_name}^{\text{dp}} \text{ some xsd:string}) \text{ and } (\text{country_name}^{\text{dp}} \text{ max } 1)) \text{ or } ((\text{locality}^{\text{dp}} \text{ some xsd:string}) \text{ and } (\text{locality}^{\text{dp}} \text{ max } 1)) \text{ or } ((\text{postal_code}^{\text{dp}} \text{ some xsd:string}) \text{ and } (\text{postal_code}^{\text{dp}} \text{ max } 1)) \text{ or } ((\text{region}^{\text{dp}} \text{ some xsd:string}) \text{ and } (\text{region}^{\text{dp}} \text{ max } 1)) \text{ or } ((\text{street_address}^{\text{dp}} \text{ some xsd:string}) \text{ and } (\text{street_address}^{\text{dp}} \text{ max } 1))$

is in range of
[has address^{dp}](#)

Aunque en el vocabulario la dirección de entrega se llame ADR, el usuario de la información entenderá que un nombre de país o un código postal pueden interpretarse como ADR.



4. Anonimizar los datos de carácter personal, información estadística, información indirectamente identificable, e información sensible o confidencial.

Anonimizar es el proceso que impide la identificación de las unidades de estudio que son fuente para los registros individuales de un conjunto⁸ de datos, o más claro, es la aplicación de técnicas que permitan expresar un dato relativo a personas, eliminando la referencia a su identidad⁹.

Básicamente lo que se busca es reducir los riesgos en el uso y procesamiento de datos personales, evitar la extracción de registros que identifican a una persona, reducir la combinación de dos o más registros en una o varias bases de datos para identificar a una persona, y minimizar la deducción de datos personales partir de uno o varios datos publicados.

¿Qué información debo prioritariamente anonimizar?

- ✓ **Datos personales:** Información numérica, alfabética, gráfica, fotográfica, acústica o cualquier otro tipo, susceptible de recogida, registro, tratamiento y transmisión, concerniente a personas físicas identificadas o identificables (tales como nombre, apellidos, estado civil, sexo, edad, domicilio, identificación personal, número de teléfono, etc.)¹⁰
- ✓ **Información sensible:** Información considerada como estrictamente confidencial. Información y características referentes a la edad, procedencia, salud, raza, religión, ideología, afiliación, finanzas, etc., se consideran de carácter sensible y requieren de una protección especial.¹¹



¿Qué logro?

- ✓ Garantizar que mis datos publicados cuentan con medidas de protección de datos personales.
- ✓ Eliminar de manera efectiva el riesgo de re-identificación

⁸ https://www.dane.gov.co/files/sen/lineamientos/DSO_020_LIN_08.pdf

⁹ <http://dle.rae.es/?id=2jjMiRi>

¹⁰ https://www.dane.gov.co/files/sen/lineamientos/DSO_020_LIN_08.pdf

¹¹ Ibidem.



- ✓ Evitar que el procesamiento de mis datos tenga consecuencias negativas para alguna persona.
- ✓ Disponer muchos más datos sin invadir la información personal.



¡Recomendaciones!

- ✓ Implemente actividades de protección de la privacidad de los datos personales antes de publicarlos.
- ✓ Elimine cualquier identificador directo de la persona (nombre, documento de identidad, etc.)
- ✓ Reemplace los identificadores directos por códigos de referencia, pseudónimos (algoritmos AES12 o un hash cifrado HMAC13), códigos arbitrarios o aleatorios.
- ✓ Si reemplazando el identificador directo aún es fácil identificar a la persona, elimine el dato.
- ✓ Clasifique y/o categorice los datos.
- ✓ Redondee los datos numéricos hacia arriba o hacia abajo.



Evalúe antes, durante y después de publicados los datos, si la anonimización de sus datos es suficiente y evalúe regularmente los riesgos existentes.



Para más información consulte los [Lineamientos para la Anonimización de micro datos](#).




Ejemplo 1:

Antes de publicar sus datos, retire el identificador directo de la persona. En este caso reemplace en número de documento que identifica al usuario y reemplácelo por un consecutivo.


¹² Advanced Encryption Standard, puede consultar como ha sido implementado este algoritmo en <https://revistas.udistrital.edu.co/ojs/index.php/Tecnura/article/view/7236/8892>

¹³ Hash-based Message Authentication Code, puede consultar como se crea este cifrado en <https://docs.oracle.com/cd/E19253-01/819-7804/wanboottasks2-51/index.html>





| Usuario | Beneficio |
|----------|--------------------------|
| 80123456 | Subsidio de vivienda |
| 51123456 | Subsidio de alimentación |
| 19123456 | Subsidio al adulto mayor |
| 53123456 | Subsidio de desempleo |




| Usuario | Beneficio |
|-----------|--------------------------|
| Usuario 1 | Subsidio de vivienda |
| Usuario 2 | Subsidio de alimentación |
| Usuario 3 | Subsidio al adulto mayor |
| Usuario 4 | Subsidio de desempleo |



Ejemplo 2:

Clasifique o categorice sus datos. En este ejemplo, asigne rangos a la edad de manera tal que el titular del registro no sea identificable por este dato.



| Usuario | Edad |
|-----------|---------|
| Usuario 1 | 11 años |
| Usuario 2 | 22 años |
| Usuario 3 | 36 años |
| Usuario 4 | 63 años |



| Usuario | Rango de edad |
|-----------|--------------------|
| Usuario 1 | Entre 10 y 15 años |
| Usuario 2 | Entre 20 y 25 años |
| Usuario 3 | Entre 35 y 40 años |
| Usuario 4 | Entre 60 y 65 años |



5. Publicar los datos en formatos estándar, reutilizables y legibles por aplicaciones informáticas.

Un formato puede ser definido como la estructura *de un disco dividido en campos y pistas según un determinado sistema operativo, lo que permite almacenar en él información*¹⁴, básicamente, es la forma en que se codifican los datos en un archivo informático.

Actualmente conocemos formatos comunes de almacenamiento como **.doc**, una extensión de Word que almacena texto y cuyo propietario es la compañía Microsoft. Pero ¿qué pasa si un usuario quiere abrir un documento .doc en un procesador de texto diferente a Word?, difícilmente va a poder conocer el contenido de dicho documento sin tener esta licencia.

Lo que queremos lograr con la implementación de esta práctica, es que todos los usuarios que encuentren mis datos, puedan accederlos sin necesidad de contar con un software o aplicación informática específica.



¿Qué logro?

- ✓ Que mis datos sean consultados, utilizados y reutilizados por más usuarios.
- ✓ Que mis datos sean más interoperables y más procesados.
- ✓ Que mis datos puedan ser extraídos, transformados y procesados por aplicaciones informáticas.
- ✓ Que mis datos puedan ser compartidos e intercambiados entre sistemas.
- ✓ Disminuir los costos y los tiempos en el uso y la transformación de mis datos.



¡Recomendaciones!

- ✓ Prefiera publicar su conjunto de datos en formatos abiertos o no propietario, ya que estos facilitan el uso, re-uso y distribución de estos. Se busca que un mayor número de aplicaciones informáticas pueda extraer, leer y transformar los datos generando mayor uso y re-uso por los usuarios.


¹⁴ <http://dle.rae.es/srv/search?m=30&w=formato>



- ✓ Disponga sus datos en múltiples formatos con el fin de incrementar el acceso a los mismos. Se busca que el mayor número de usuarios pueda utilizar los datos sin tener que transformarlos primero en su formato preferido.
- ✓ Utilice Web Services para mejorar el intercambio e interoperabilidad de los datos publicados.



Ejemplo 1:

A continuación, encontrará extensiones de formato propietario y no propietario, prefiera para la publicación de sus datos las extensiones de la derecha. 

| FORMATO PROPIETARIO | |
|---------------------|-----|
| DOC | GDB |
| XLS | DWG |
| PPT | SHP |

| FORMATO NO PROPIETARIO | | | |
|------------------------|------------|-----|-----|
| TXT | ASCII Grid | KML | RDF |
| CSV | GeoJSON | DXF | XML |
| JSON | GML | ODF | |



Ejemplo 2:

Si no fue posible utilizar una extensión de formato no propietario, publique sus datos en una variedad de formatos que le permita al usuario acceder a los datos desde diferentes aplicaciones informáticas.





Ejemplo 3:

Para publicar sus datos, conozca diferentes tipos de Web Services que viajan en la red utilizando extensiones XML:

| | |
|------|---|
| SOAP | https://www.w3.org/TR/soap/ |
| REST | http://rest.elkstein.org/ |
| WMS | http://www.opengeospatial.org/standards/wms |
| WFS | http://www.opengeospatial.org/standards/wfs |
| WCS | http://www.opengeospatial.org/standards/wcs |



6. Proporcionar un diccionario de datos que facilite el uso e interpretación de los datos publicados.

Un diccionario está definido como un *repertorio en forma de libro o en soporte electrónico en el que se recogen, según un orden determinado, las palabras o expresiones de una o más lenguas, o de una materia concreta, acompañadas de su definición, equivalencia o explicación*¹⁵. específicamente, en un diccionario de datos se encuentran los conceptos de los datos y sus propiedades (atributos, dominios, asociaciones y operaciones), acompañados de una definición en donde se identifica el propósito, alcance, y campo de aplicación del dato publicado.

Lograr que los usuarios de mis datos entiendan la información que está publicada, le permitirá decidir si satisfacen o no su necesidad, y le facilitará la decisión para explotar lo datos.



Para más información consulte el [Instructivo Diccionario de Datos](#).



¿Qué logro?

- ✓ Que mis datos sean consultados, utilizados y reutilizados por más usuarios.
- ✓ Que se eviten ambigüedades en la interpretación de mis datos.
- ✓ Que mis datos publicados estén unificados y consistentes.
- ✓ Que los usuarios comprendan mejor y dispongan de claridad sobre los conceptos de mis datos.



¡Recomendaciones!

- ✓ Publique como mínimo un diccionario de datos que contenga la caracterización del conjunto de datos producidos por su entidad.
- ✓ Cree el diccionario de datos teniendo en cuenta como mínimo la siguiente información:

| | |
|----------------------------|---|
| Nombre del atributo | Nombre que identifica de manera clara un atributo |
| Definición | Descripción del atributo |

¹⁵ <http://dle.rae.es/?w=diccionario>



| | |
|------------------|--|
| Tipo de dato | Tipo de valor del atributo |
| Unidad de medida | Expresa la unidad de medida del atributo |
| Dominio | Determina si el atributo tiene asociado un dominio |



Ejemplo:

Para el conjunto de datos “Establecimientos Educativos en Bogotá”, publicaríamos un diccionario de datos similar al siguiente:

| Nombre del atributo | Definición | Tipo de dato | Unidad de Medida | Dominio |
|-------------------------------------|--|----------------------|------------------|----------|
| Código del establecimiento | Código DANE asignado al Institución Educativa | Decimal | | |
| Nombre del establecimiento | Nombre de la Institución Educativa | Cadena de Caracteres | | |
| Área construida del establecimiento | Área construida de la Institución Educativa | Decimal | Metros cuadrados | |
| Sector | Clasificación asignada según la naturaleza jurídica, la cual define las principales características que desde lo jurídico - administrativo y tiene que ver con el origen de su creación. | Entero | | EESector |

No podríamos olvidar detallar el dominio asociado,

| Atributo | Nombre del Dominio | Código | Definición |
|----------|--------------------|--------|---|
| EESector | Oficial | 1 | Es una institución donde su naturaleza de servicio público cultural y su función social le es inherente, cumpliendo con las disposiciones legales y estatutarias que las rigen, donde sus rentas se conservan y se aplican debidamente. |
| | Privado | 2 | Es una institución, constituida como persona jurídica de utilidad común, donde sus rentas se conservan y se aplican debidamente y que en todo lo esencial se cumple con la voluntad de sus fundadores. |



7. Acompañar la publicación de los datos con enlaces a otros recursos.

Un enlace es definido como una secuencia de caracteres que se utiliza como dirección para acceder a información adicional¹⁶, más sencillo puede definirse como una expresión que conecta una cierta información con otra¹⁷. Ahora bien, un recurso se puede entender como cualquier medio que satisfaga una necesidad, por ejemplo: datos, servicios, documentos, o software.

En términos generales, se sugiere que para facilitar al usuario la comprensión de los datos publicados, se proporcionen recursos o documentos adicionales que puedan ser consultados y/o descargados través de enlaces o links de otro sitio Web.

Dentro de los recursos adicionales que pueden publicarse para acompañar el dato, están: los procedimientos de captura del dato, los catálogos de objetos, las especificaciones técnicas, los documentos normativos, los catálogos de simbología, los reportes de evaluación de calidad, entre otros.



¿Qué logro?

- ✓ Mejorar la consulta, utilización y reutilización de mis datos.
- ✓ Que los usuarios comprendan mejor mis datos
- ✓ Que mis datos sean compartidos e intercambiados por un mayor número de usuarios.



¡Recomendaciones!

- ✓ Disponga información adicional de sus datos que le permita a los usuarios ampliar el horizonte de las aplicaciones y funcionalidades en las que pueda utilizar y explotar del dato.



No olvide vincular o registrar esta información en el metadato.

¹⁶ <http://dle.rae.es/?id=FTVRKvj>





¹⁷ <https://sistemas.com/enlace.php>





Ejemplo 1:

El “Mapa de Referencia para Bogotá” publicado en <https://www.ideca.gov.co/datos-de-referencia>, tiene publicados varios enlaces a otros recursos:

-  [Especificación técnica para la elaboración del Mapa de Referencia para Bogotá D.C. V.03.18](#)
-  [Catálogo de Objetos del Mapa de Referencia de Bogotá D.C V 5.7](#)
-  [Catálogo de Representación del Mapa de Referencia de Bogotá D.C. V.3.0](#)
-  [Reporte de evaluación de calidad generado al Mapa de Referencia del Distrito Capital V.12.17](#)



Ejemplo 2:

La información de los “humedales de Bogotá” publicada en <http://ambientebogota.gov.co/web/sda/humedales>, tiene publicados enlaces a otros recursos:

- [Gestión de los humedales](#)
- [Política de Humedales del Distrito Capital](#)
- [Normatividad](#)
- [Planes de Manejo Ambiental](#)
- [Comité distrital de humedales, documentos de consulta normatividad](#)



8. Depurar sus datos antes de publicarlos para garantizar su calidad y aumentar su confianza.

La calidad está definida como *el conjunto de propiedades inherentes a algo, que permiten juzgar su valor¹⁸*, para el caso que nos concierne, este algo hace referencia al dato publicado. Ahora bien, la calidad también es entendida como la capacidad del dato publicado para satisfacer los requerimientos y las necesidades de los usuarios, respetando una serie de normas y especificaciones establecidas.

Actualmente, existen múltiples elementos de calidad de los datos que identifican las cualidades de estos, sin embargo, en este aparte citaremos los elementos principales que usted o su entidad deben tener en cuenta para publicar sus datos con un estándar mínimo de calidad.



Para más información consulte el [Procedimiento para evaluar y reportar la calidad de los datos espaciales](#).



¿Qué logro?

- ✓ Que mis datos sean consultados, utilizados y reutilizados por más usuarios.
- ✓ Que los usuarios hablen bien de mis datos.
- ✓ Que mis datos sean más interoperables y más procesados.
- ✓ Que mis datos sean más compartidos.
- ✓ Que mis datos sean más confiables.



Recomendación 1:

Establezca si dentro de su conjunto de datos existen capos vacíos o nulos, con información descontextualizada, sobrante, o duplicada. Si existen, complételos o elimínelos según corresponda antes de la publicación. El elemento de calidad asociado se denomina *“totalidad”*.

¹⁸ <http://dle.rae.es/?id=6nVpk8P|6nXVL1Z>



| | ❌ | ✅ | | | | | | | | | | | | | | | | | | |
|---------------------|--|-----------|-----------|------|-----|------|-----|----------|-----|--|-----|------------|-----|--|-----------|-----------|------|-----|------------|-----|
| | <table border="1"> <thead> <tr> <th>Localidad</th> <th>Atendidos</th> </tr> </thead> <tbody> <tr> <td>Suba</td> <td>200</td> </tr> <tr> <td>Suba</td> <td>200</td> </tr> <tr> <td>Américas</td> <td>100</td> </tr> <tr> <td></td> <td>850</td> </tr> <tr> <td>Tunjuelito</td> <td>500</td> </tr> </tbody> </table> | Localidad | Atendidos | Suba | 200 | Suba | 200 | Américas | 100 | | 850 | Tunjuelito | 500 | <table border="1"> <thead> <tr> <th>Localidad</th> <th>Atendidos</th> </tr> </thead> <tbody> <tr> <td>Suba</td> <td>200</td> </tr> <tr> <td>Tunjuelito</td> <td>500</td> </tr> </tbody> </table> | Localidad | Atendidos | Suba | 200 | Tunjuelito | 500 |
| Localidad | Atendidos | | | | | | | | | | | | | | | | | | | |
| Suba | 200 | | | | | | | | | | | | | | | | | | | |
| Suba | 200 | | | | | | | | | | | | | | | | | | | |
| Américas | 100 | | | | | | | | | | | | | | | | | | | |
| | 850 | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500 | | | | | | | | | | | | | | | | | | | |
| Localidad | Atendidos | | | | | | | | | | | | | | | | | | | |
| Suba | 200 | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500 | | | | | | | | | | | | | | | | | | | |
| Campo duplicado | → | | | | | | | | | | | | | | | | | | | |
| Campo inconsistente | → | | | | | | | | | | | | | | | | | | | |
| Campo vacío | → | | | | | | | | | | | | | | | | | | | |

📌 Recomendación 2:

Identifique y transforme si es necesario, los datos que hayan sido registrados en un formato diferente al establecido para su conjunto de datos. Es decir, estandarice el formato de salida del dato que va a publicar. El elemento de calidad asociado se denomina “*consistencia de formato*”.

| | ❌ | | ✅ | | | | | | | | | | | | | | | | | | | | |
|------------|---|-----------|-----------|------|-----|------------|-------|----------|------|------------|-----|-------------------------|--|-----------|-----------|------|-----|------------|-----|----------|-----|------------|-----|
| | <table border="1"> <thead> <tr> <th>Localidad</th> <th>Atendidos</th> </tr> </thead> <tbody> <tr> <td>Suba</td> <td>200</td> </tr> <tr> <td>Tunjuelito</td> <td>500,8</td> </tr> <tr> <td>Fontibón</td> <td>40 %</td> </tr> <tr> <td>Tunjuelito</td> <td>500</td> </tr> </tbody> </table> | Localidad | Atendidos | Suba | 200 | Tunjuelito | 500,8 | Fontibón | 40 % | Tunjuelito | 500 | } Formato inconsistente | <table border="1"> <thead> <tr> <th>Localidad</th> <th>Atendidos</th> </tr> </thead> <tbody> <tr> <td>Suba</td> <td>200</td> </tr> <tr> <td>Tunjuelito</td> <td>500</td> </tr> <tr> <td>Fontibón</td> <td>430</td> </tr> <tr> <td>Tunjuelito</td> <td>500</td> </tr> </tbody> </table> | Localidad | Atendidos | Suba | 200 | Tunjuelito | 500 | Fontibón | 430 | Tunjuelito | 500 |
| Localidad | Atendidos | | | | | | | | | | | | | | | | | | | | | | |
| Suba | 200 | | | | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500,8 | | | | | | | | | | | | | | | | | | | | | | |
| Fontibón | 40 % | | | | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500 | | | | | | | | | | | | | | | | | | | | | | |
| Localidad | Atendidos | | | | | | | | | | | | | | | | | | | | | | |
| Suba | 200 | | | | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500 | | | | | | | | | | | | | | | | | | | | | | |
| Fontibón | 430 | | | | | | | | | | | | | | | | | | | | | | |
| Tunjuelito | 500 | | | | | | | | | | | | | | | | | | | | | | |

📌 Recomendación 3:

Si sus datos tienen algún tipo de clasificación o rangos, previo a la publicación valide que estos se hayan registrado correctamente. El elemento de calidad asociado se denomina “*exactitud de clasificación*”.

| | ❌ | | ✅ | | | | | | | | | | | | | | | | | | | | | | | | |
|-----------|---|----------------|------|---------------|----------|---|---------------|-----------|----|---------------|----------|----|----------------|-------------------------------|--|--------|------|---------------|----------|---|---------------|-----------|----|----------------|----------|----|----------------|
| | <table border="1"> <thead> <tr> <th>Género</th> <th>Edad</th> <th>Rango de Edad</th> </tr> </thead> <tbody> <tr> <td>Femenino</td> <td>5</td> <td>0 < Edad ≤ 10</td> </tr> <tr> <td>Masculino</td> <td>15</td> <td>0 < Edad ≤ 10</td> </tr> <tr> <td>Femenino</td> <td>35</td> <td>40 < Edad ≤ 50</td> </tr> </tbody> </table> | Género | Edad | Rango de Edad | Femenino | 5 | 0 < Edad ≤ 10 | Masculino | 15 | 0 < Edad ≤ 10 | Femenino | 35 | 40 < Edad ≤ 50 | } Clasificación inconsistente | <table border="1"> <thead> <tr> <th>Género</th> <th>Edad</th> <th>Rango de Edad</th> </tr> </thead> <tbody> <tr> <td>Femenino</td> <td>5</td> <td>0 < Edad ≤ 10</td> </tr> <tr> <td>Masculino</td> <td>15</td> <td>10 < Edad ≤ 20</td> </tr> <tr> <td>Femenino</td> <td>35</td> <td>30 < Edad ≤ 40</td> </tr> </tbody> </table> | Género | Edad | Rango de Edad | Femenino | 5 | 0 < Edad ≤ 10 | Masculino | 15 | 10 < Edad ≤ 20 | Femenino | 35 | 30 < Edad ≤ 40 |
| Género | Edad | Rango de Edad | | | | | | | | | | | | | | | | | | | | | | | | | |
| Femenino | 5 | 0 < Edad ≤ 10 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Masculino | 15 | 0 < Edad ≤ 10 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Femenino | 35 | 40 < Edad ≤ 50 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Género | Edad | Rango de Edad | | | | | | | | | | | | | | | | | | | | | | | | | |
| Femenino | 5 | 0 < Edad ≤ 10 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Masculino | 15 | 10 < Edad ≤ 20 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Femenino | 35 | 30 < Edad ≤ 40 | | | | | | | | | | | | | | | | | | | | | | | | | |





Recomendación 4:

Si sus datos representan atributos cuantitativos, verifique antes de la publicación que estos correspondan a valores conocidos como verdaderos. El elemento de calidad asociado se denomina “*exactitud de atributos cuantitativos*”.

| ✘ | | | ✔ | | |
|------------|--------|-------------|------------|--------|-------------|
| Localidad | Mes | Temperatura | Localidad | Mes | Temperatura |
| Chapinero | Enero | 7°C | Chapinero | Enero | 7°C |
| Suba | Agosto | -15°C | Suba | Agosto | 15°C |
| Tunjuelito | Abril | 60°C | Tunjuelito | Abril | 16°C |

} Valor inconsistente



Recomendación 5:

Si los datos a publicar tienen asociadas fechas de captura o recolección de información (Por ejemplo, registro de incidentes, registro de vacunación, conteo de vehículos, etc.), valide que las fechas registradas coincidan con la fecha del evento capturado. El elemento de calidad asociado se denomina “*exactitud de la medición del tiempo*”.

| ✘ | | ✔ | |
|--------------------|-------------------|--------------------|-------------------|
| Precipitación (mm) | Fecha de registro | Precipitación (mm) | Fecha de registro |
| 0,0 | 20580101 | 0,0 | 20180101 |
| 0,1 | 20180102 | 0,1 | 20180102 |
| 12,6 | 20510403 | 12,6 | 20180103 |
| 2,0 | 20180104 | 2,0 | 20180104 |

← Fecha inconsistente



Recomendación 6:



UAECD

Cuando sus datos reflejen información determinada por un orden de eventos, garantice antes de la publicación que este orden sea el adecuado. El elemento de calidad asociado se denomina “consistencia temporal”.

| ❌ | | | } Fechas no secuenciales | ✅ | | |
|-----------|-------------|-------------------------|--------------------------|------------|-------------------------|--------------------|
| Usuario | Escolaridad | Periodo de Estudio | | Usuario | Escolaridad | Periodo de Estudio |
| Usuario 1 | Preescolar | 2003/01/01 - 2005/12/31 | Usuario 1 | Preescolar | 2003/01/01 - 2005/12/31 | |
| Usuario 1 | Primaria | 2003/01/01 - 2010/12/31 | Usuario 1 | Primaria | 2006/01/01 - 2010/12/31 | |
| Usuario 1 | Secundaria | 2006/01/01 - 2016/31/12 | Usuario 1 | Secundaria | 2011/01/01 - 2014/31/12 | |
| Usuario 1 | Media | 2015/01/01 - 2016/31/12 | Usuario 1 | Media | 2015/01/01 - 2016/31/12 | |

 **Recomendación 7:**

Si sus datos incluyen alguna ubicación (dirección, lugar, coordenadas), verifique que esta corresponda a la ubicación real. El elemento de calidad asociado se denomina “exactitud posicional”.

| ❌ | | ← Lugar inexistente | ✅ | |
|--------------------|----------------------|---------------------|----------------------|-------|
| Eventos Deportivos | Lugar | | Eventos Deportivos | Lugar |
| Evento 1 | AK 30 25 90 | Evento 1 | AK 30 25 90 | |
| Evento 2 | AC 63 215 20 | Evento 2 | Parque La Florida | |
| Evento 3 | Parque Simón Bolívar | Evento 3 | Parque Simón Bolívar | |



9. Incorporar el componente geográfico en los datos publicados para optimizar su comprensión, uso, re-uso y visualización.

El espacio geográfico es un punto de la superficie terrestre que se vea afectado o que afecte de alguna manera a la humanidad¹⁹. Este punto de la superficie puede ser la dirección de una institución educativa, la ubicación de un centro de vacunación, el lugar donde se miden las variaciones de tránsito en la ciudad, los parques de Bogotá, el edificio en donde se llevan a cabo los trámites de una entidad, etc.

Los datos que usted captura o recolecta en su quehacer misional siempre están asociados a un espacio geográfico, por lo tanto, este componente debería ser incluido en los datos publicados. Tenga en cuenta que el componente geográfico lo puede incorporar incluyendo una ubicación asociada a una dirección, a un lugar, a una localidad, a un barrio, al nombre de un establecimiento públicamente reconocido, o a cualquier punto que permita al usuario identificar la localización del dato publicado.



¿Qué logro?

- ✓ Ampliar el espectro de usuarios que consultan, utilizan y reutilizan mis datos.
- ✓ Que mis datos sean más interoperables y más procesados.
- ✓ Que mis datos sean puedan compartir en más plataformas.
- ✓ Que los usuarios puedan visualizar mis datos en un mapa.



¡Recomendaciones!

Disponga sus datos agregando un atributo de ubicación como direcciones, coordenadas o lugares geográficos.

¹⁹ http://enciclopedia.banrepcultural.org/index.php?title=Espacio_geogr%C3%A1fico





Ejemplo 1:

Incluya direcciones que permitan localizar espacialmente sus datos.

| Programa | Dirección |
|---|---------------------|
| Registro de atención a la ciudadanía | CL 138 108 16 |
| Centro de autocuidado | KR 18 D 77 A 11 SUR |
| Familias con un niño en cupos cofinanciados | CL 70 A SUR 20 46 |
| Centro de autocuidado | KR 18 D 73 A 33 SUR |



Ejemplo 2:

Utilice coordenadas para indicar la ubicación de sus datos.

| EPS | Coordenada X | Coordenada Y |
|-------------------|--------------|--------------|
| EPS Convida | 90874.96334 | 105134.2164 |
| Capital Salud EPS | 98632.1006 | 95487.5808 |
| EPS Famisanar | 97425.08812 | 88370.31501 |
| Aliansalud EPS | 97169.43356 | 96597.06891 |
| Salud Total EPS | 92156.24975 | 96251.89518 |



Tenga en cuenta que, para publicar datos con coordenadas, tendrá que incluir dentro de su metadato el Sistema de referencia empleado

MAGNA – SIRGAS EPSG: 4686 <https://epsg.io/4686>

MAGNA – SIRGAS EPSG: 3116 <https://epsg.io/3116>

MAGNA – BOGOTÁ SR-ORG: 7458 <http://spatialreference.org/ref/sr-org/7458/>





Ejemplo 3:

También puede incluir lugares geográficos descriptivos así:

| Usuario Vacunado | Lugar de Vacunación |
|------------------|---------------------|
| Usuario 1 | BOSQUE IZQUIERDO |
| Usuario 2 | LAGUNITAS URBANO |
| Usuario 3 | LA CAPUCHINA |
| Usuario 4 | SAN JOSE SUR |
| Usuario 5 | VENECIA |



10. Asignar un identificador único que permita aumentar la consulta, uso e interoperabilidad del dato publicado.

Considere un identificador único como un texto que permite identificar inequívocamente el dato publicado. Si usted va a publicar un dato buscando que este sea reutilizado por los usuarios y por su propia entidad, asignar un identificador facilitará el descubrimiento y la citación del dato a través del tiempo, independientemente del estado, la disponibilidad o el formato de este.

En la vida cotidiana, mediante la cédula de ciudadanía se asigna a cada ciudadano un número de identificación único, y a través de este identificador el ciudadano es reconocido a lo largo de su vida e incluso después de su muerte por otras personas y/o autoridades. En el mundo de la información, cada dato debería viajar con un identificador único que le permita ser encontrado y reconocido fácilmente entre los demás datos.

Ahora bien, si usted o su entidad publican continuamente datos, establecer un sistema de identificación común le permitirá identificar rápidamente los datos y los hará comparables por cualquier usuario de manera confiable. Esta condición es previa para lograr una correcta gestión y reutilización de sus datos.



¿Qué logro?

- ✓ Ampliar el espectro de usuarios que consultan, utilizan y reutilizan mis datos.
- ✓ Que mis datos sean más interoperables y más procesados.
- ✓ Que mis datos sean compartidos en más plataformas.
- ✓ Que mis datos puedan vincularse a otros datos en el entorno Web.
- ✓ Que mis datos sean encontrados, usados y citados inequívocamente.
- ✓ Garantizar y monitorear la reutilización de mis datos.
- ✓ Reducir del uso erróneo de mis datos.



¡Recomendaciones!

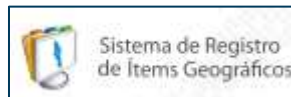
- ✓ Establezca un sistema de identificación común para los datos que va a publicar.
- ✓ Utilice siempre el identificador asignado para publicar sus datos.





Ejemplo:

IDECA ha dispuesto el Sistema de Registro de Ítems geográficos <http://geoitems.ideca.gov.co/geoitems/ano/> , una herramienta web desarrollada para la creación, aprobación, publicación y mantenimiento de ítems geográficos que permite la asignación concertada de identificadores y significados únicos que contribuyen a la definición de un lenguaje común para el desarrollo de nuevos sistema de información en el Distrito, garantizando con ello su fiabilidad e interoperabilidad. Adicionalmente, permite almacenar descripciones de la información oficial producida e identificar su respectivo custodio jugando así un papel importante en el control y seguimiento de la producción de información geográfica.



Este sistema asigna un identificador único a cada uno de los ítems creados. Dentro de los ítems publicados se reconoce el **ID 65** como unidad de planeamiento, y siempre que se hace referencia a este objeto es suficiente nombrarlo así para entender que se habla de “Unidades territoriales de análisis, planeamiento y gestión que tienen como propósito definir y precisar el planeamiento del suelo urbano, de expansión y rural”.

Búsqueda

Nombre del ítem:

[Búsqueda avanzada](#)

7 resultados encontrados

| ID | Nombre del ítem | Custodio | Clase de ítem | Estado | Fecha de aceptación | Fecha de modificación |
|-----|---|---|---------------|--------|---------------------|-----------------------|
| 31 | Unidad administrativa | Secretaría Distrital de Planeación - SDP | Grupo | Válido | 2012-05-28 | 2012-05-28 |
| 65 | Unidad de planeamiento | Secretaría Distrital de Planeación - SDP | Objeto | Válido | 2012-05-28 | 2012-05-28 |
| 37 | Unidad de planeamiento zonal | Secretaría Distrital de Planeación - SDP | Subtipo | Válido | 2012-05-28 | 2012-05-28 |
| 40 | Unidad de planeamiento rural | Secretaría Distrital de Planeación - SDP | Subtipo | Válido | 2012-05-28 | 2012-05-28 |
| 304 | Unidad de Protección Integral | Instituto para la Protección de la Infancia y la Juventud - IDIRIOJ | Objeto | Válido | 2013-12-24 | 2013-12-24 |
| 478 | Unidad de Mediación y Concilia | Secretaría Distrital de Gobierno - SDG | Objeto | Válido | 2013-05-08 | 2013-05-08 |
| 32 | Unidad territorial y unidad adm | Secretaría Distrital de Planeación - SDP | Tema | Válido | 2012-05-28 | 2012-05-28 |



Obtenga el más Alto Nivel en Publicación de Datos

Hasta aquí hemos publicado las que consideramos como mejores prácticas de publicación de datos. Ahora, usted como publicador o como entidad debe evaluar si sus datos se están disponiendo con los más altos y mejores estándares de publicación.

Lo invitamos a que valide en qué nivel de publicación se encuentra su entidad teniendo en cuenta el siguiente listado:



Nivel Básico

En este nivel se encuentran aquellas entidades que publican sus datos en formatos estándar y reutilizables ([práctica #5](#)), que integran a la publicación sus respectivos metadatos ([práctica #1](#)), y que acompañan la publicación con una licencia asociada ([práctica #2](#)).



Nivel Intermedio

Este nivel lo alcanzan las entidades que superando el nivel básico disponen un mayor volumen de información gracias a que anonimizan sus datos ([práctica #4](#)), implementan medidas de calidad que aumentan la confianza en sus datos ([práctica #8](#)), incorporan el componente geográfico proporcionando valor agregado a sus datos ([práctica #9](#)), y publican un diccionario de datos que facilita la interpretación de los datos publicados ([práctica #6](#)).



Nivel Avanzado

Las entidades que superando el nivel intermedio utilizan vocabularios estandarizados en sus nuevas publicaciones ([práctica #3](#)), acompañan la publicación de sus datos con enlaces a otros recursos ([práctica #7](#)), y asignan un identificador único para dato logrando así aumentar su descubrimiento y reutilización ([práctica 10](#)).



Referencias

Bernadette Farias Lóscio, Caroline Burle, Newton Calegari. W3C. Data on the Web Best Practices. 31 de enero de 2017. W3C Recommendation. URL: <https://www.w3.org/TR/dwbp/>

Bernadette Hyland; Ghislain Auguste Atemezang; Boris Villazón-Terrazas. W3C. Best Practices for Publishing Linked Data. 9 de enero de 2014. W3C Note. URL: <https://www.w3.org/TR/ld-bp/>

Dan Brickley; Libby Miller. FOAF project. FOAF Vocabulary Specification 0.99 (Paddington Edition). 14 de enero de 2014. URL: <http://xmlns.com/foaf/spec/>

Dublin Core metadata initiative. DCMI Metadata Terms. 14 de junio de 2012. DCMI Recommendation. URL: <http://dublincore.org/documents/dcmi-terms/>

Fadi Maali; John Erickson. W3C. Data Catalog Vocabulary (DCAT). 16 de enero de 2014. W3C Recommendation. URL: <https://www.w3.org/TR/vocab-dcat/>

ISO 19115-1:2014 (E). Geographic information – Metadata. Part1: Fundamentals. 01 de abril de 2014. International Standard.

ISO 19126:2009 (E). Geographic information -- Feature concept dictionaries and registers. noviembre de 2009. International Standard.

ISO 19157:2013 (E). Geographic information – Data quality. 15 de diciembre de 2013. International Standard.

Jeremy Tandy; Payam Barnaghi; Linda van den Brink. W3C. Spatial Data on the Web Best Practices. 5 de enero 2017. W3C Note. URL: <https://www.w3.org/TR/sdw-bp/>

T. Berners-Lee; R. Fielding; L. Masinter. IETF. Uniform Resource Identifier (URI): Generic Syntax. Enero de 2005. Internet Standard. URL: <https://tools.ietf.org/html/rfc3986>

World Wide Web Foundation. Open Data Barometer. URL: <http://opendatabarometer.org>



W3C OWL Working Group. W3C. OWL 2 Web Ontology Language Document Overview (Second Edition). 11 de diciembre de 2012. W3C Recommendation. URL: <https://www.w3.org/TR/owl2-overview/>

W3.org. Semantic Web - W3C. URL: <https://www.w3.org/standards/semanticweb/> Consultado el 12 de septiembre de 2018.





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

Latitud: 4.603557, Longitud: -74.094105
Bogotá, Cundinamarca, Colombia.