

# La evaluación de calidad de los metadatos geográficos como un aporte para la consolidación de la ICDE

Jorge Eliécer Arias Rincón<sup>1</sup>, Luis David Garnica Olarte<sup>2</sup>, Fredy Montealegre<sup>3</sup>

## Resumen

El metadato geográfico es un recurso valioso a la hora de consultar, divulgar, acceder y distribuir los datos espaciales<sup>4</sup> o la información geográfica<sup>5</sup>; se ha convertido en base fundamental para las infraestructuras de datos espaciales (en adelante IDE) vistas estas según Granell. C. y Gould. M. (2006) como sistemas basados en la apertura de servicios estandarizados, accesibles a través de la red, que proporcionan una infraestructura libre y generalista que tienden a la máxima difusión aprovechable por todo tipo de usuarios; de ahí la importancia de contar con información confiable que haga referencia a los datos espaciales. Un producto cartográfico o un dato espacial o la información geográfica pierden sentido cuando esta no es utilizada, cuando es tergiversada desde su origen. Por eso, en la búsqueda de prevenir que se presenten errores en el trabajo de producción, publicación y difusión de la información se presenta un método que ha sido de gran ayuda a la hora de realizar el control de calidad de los metadatos geográficos, generados en los diversos proyectos que ha liderado la Infraestructura Colombiana de Datos Espaciales (ICDE).

## Palabras claves

*Calidad de información, técnica de muestreo, metadato geográfico, información geográfica, datos espaciales, plan de muestreo, Infraestructura de Datos Espaciales (IDE), Infraestructura Colombiana de Datos Espaciales (ICDE).*

## Abstract

The geographic metadata is a valuable resource when it comes to access, disseminate and distribute data access spatial or geographical information. It has become the foundation for Spatial Data Infrastructures (the SDI) views these as Granell. C. and Gould. M. (2006) as systems based on open standardized services accessible through the network to provide free and comprehensive infrastructure which tend to the maximum diffusion usable by all users. Therefore the importance of reliable information that refers to the spatial data. It is fundamental to the potential success of an SDI, a data mapping product or a spatial or geographic information lose meaning when it is not used when it is distorted from its origin. By that in seeking to prevent errors occur in this type that would throw overboard all the work of production, publication and dissemination of information. A method has been of great help in performing the quality control of geographic metadata generated in the various projects he has led the Colombian Spatial Data Infrastructure (hereinafter CSDI).

## Key words

*Information quality, sampling technique, geographic metadata, geographic information, spatial data, sampling plan, Spatial Data Infrastructure, SDI, Colombian Spatial Data Infrastructure, CSDI.*

- 1 Ingeniero Catastral y Geodesta de la Universidad Francisco José de Caldas, estudiante de la maestría en Geografía, Universidad Pedagógica y Tecnológica de Colombia (UPTC). Investigador en el tema de estándares de información geográfica para el grupo de Infraestructura de Datos Espaciales (IDE), del Centro de Investigación y Desarrollo de Información Geográfica (CIAF), del Instituto Geográfico Agustín Codazzi (IGAC).
- 2 Ingeniero Forestal, Especialista en gerencia de Proyectos – Instituto Geográfico Agustín Codazzi, Subdirección de Agrología - Gurpo de trabajo de Cobertura y Uso de la Tierra, Bogotá D.C., Colombia.dgarnica@igac.gov.co,david.forestal@gmail.com
- 3 Ingeniero Geógrafo, especialista en Sistemas de Información Geográfica, candidato a magíster en Sistemas de Información Geográfica, Instituto Geográfico Agustín Codazzi, Oficina Centro de Investigación y Desarrollo en Información Geográfica - CIAF, Bogotá D.C., Colombia. fredy.montealegre@igac.gov.co;fredy.montealegre@gmail.com
- 4 Un dato espacial es la representación de la realidad geográfica, la cual es concebida como la suma de objetos que cumple con tres características fundamentales: en primera medida deben estar asociados a una localización, es decir, pueden ser georeferenciados, en segundo lugar deben tener atributos que los definan y los hagan únicos, es decir, son definidos por sus cualidades ya sean estas propiedades físicas, químicas, o percepciones, y por último se pueden asociar con otros objetos a partir de sus relaciones u operaciones ya sean geométricas (topología) o establecidas a partir de las relaciones como sensación humana.
- 5 Los datos por sí solos no son más que unidades espaciales que permiten inferir su forma real en un universo de discurso; la agrupación de estos elementos u objetos conforma estructuras más complejas debido a la interacción o relación entre ellos y a su jerarquización; esto es denominado como información geográfica. Según Cebrian. J (1992) los elementos o unidades espaciales son el material de la información geográfica.



## Introducción

El documento se divide en dos partes: en la primera se establecen los pasos previos que son necesarios para garantizar un nivel mínimo de calidad, asegurando por medio del uso de plantillas y la implementación de un catálogo de metadatos que los errores de tipo lógico en la información no se presenten o se minimicen. En la segunda parte se describe el procedimiento específico utilizado para realizar el control de calidad de un conjunto de metadatos migrados de un estándar a metadatos conforme con la Norma Técnica Colombiana NTC 4611, utilizando un plan de calidad que asocia un muestreo aleatorio simple, con niveles aceptables de calidad NAC y métodos estadísticos para la determinación del error.

El procedimiento es de gran importancia, teniendo en cuenta la transición de la versión de la NTC de primera a segunda actualización. Los resultados del control de calidad dentro de los futuros operativos y la interpretación de las cifras, derivadas del cálculo de las variables<sup>6</sup> estudiadas, entregan una idea objetiva de la calidad del proceso y permiten visualizar el conjunto de datos de lo general a lo particular. Por medio de la interpretación de los datos, ya sean cualitativos o cuantitativos, se puede definir el grado de error de la muestra.

Esto se refleja en el universo de metadatos validados y permite establecer los casos críticos y realizar las correcciones para la optimización del proceso.

## 1. Usando plantillas de metadato

En el contexto del metadato aplicado en las IDE se asocia este siempre a un estándar de información geográfica para su generación. En el caso colombiano se refiere a la NTC 4611<sup>7</sup>, sin embargo, no se puede desconocer que existe gran información documentada por medio de otros estándares, pero en el desarrollo de este artículo se hará referencia a las implementaciones que se realizan en la ICDE, cuyas entidades adscritas siguen los lineamientos según el documento 3585<sup>8</sup> del 2009 del Consejo Nacional de Política Económica y Social (CONPES), el cual define que la ICDE implementará las normas generadas en el comité técnico de normalización de la información geográfica: Comité 028.

Generando metadatos a partir de la norma NTC 4611, se facilita la evaluación de calidad, esto debido a que la norma permite generar perfiles<sup>9</sup> institucionales de metadato, los cuales son subconjuntos de elementos que aplican

6 Los elementos que conforman las secciones de un metadato y que están definidas dentro de la Norma Técnica Colombiana NTC 4611.

7 Norma Técnica Colombiana NTC 4611 Metadatos Geográficos; actualmente la versión oficial para la generación de metadatos es la segunda actualización de la norma con fecha de oficialización del 13 de abril de 2011, por el Instituto Colombiano de Normas Técnicas (ICONTEC). <http://www.icontec.org.co/index.php?section=1>

8 En su apartado D, el CONPES expone: Todas las entidades del Estado y aquellas de carácter mixto o privado que ejerzan funciones públicas deberán seguir en la producción o adquisición de IG, los lineamientos y normas técnicas definidas en el marco del Comité Técnico de Normalización de la Información Geográfica 028 de ICONTEC, en el cual participan instituciones que integran la ICDE.

específicamente a la información temática que desarrolla, administra o genera una entidad. Se pueden dar casos en los cuales la información temática obedezca a características particulares para la entidad, y se dificulta documentar aquellos elementos de dato que no aplican a la información. Por esto se realiza un estudio previo de la conveniencia de los elementos a incluirse en el futuro metadato y se crea el perfil institucional.

Ya con la disposición de un perfil institucional, se pueden desprender de este subconjuntos de elementos de dato que permitan documentar información especializada. Por ejemplo, si una entidad como la Agencia Nacional de Hidrocarburos requiere documentar sus metadatos, el primer paso sería determinar cuáles son los campos temáticos que desarrolla en cuanto a producción de información; en segundo lugar, y a partir de ese diagnóstico, se debe generar la plantilla institucional, que agrupará todos los elementos de dato que en algún momento pueda utilizar para documentar cualquier producto.

Hasta el momento solo se tiene una agrupación de elementos de dato que han sido identificados por su uso potencial y se han reducido los campos que se tendrían que documentar; sin embargo, generar el metadato a partir de la plantilla institucional seguirá siendo complicado por la posible presencia de elementos que no se van a usar. La solución a este problema es agrupar la información por temas clasificándolos por sus características homogéneas, por ejemplo, para el caso de los productos asociados siempre a la infraestructura petrolera se puede generar un grupo de información que

los identifique, o para los mapas topográficos también se puede generar otro grupo, para la hidrografía, o los mapas de riesgo, y así para cualquier tipo de información. Además, dentro de estos temas se puede derivar información aún más especializada y se pueden tener como criterios de selección los niveles de representación que muestren, por ejemplo, si se tiene un grupo de imágenes, estas podrían agruparse por los sensores que las producen, o si se tiene la cartografía vectorial puede desagregarse por sus escalas.

Cuando se quiere documentar a nivel especializado la información, como producto es importante crear una plantilla de metadato, definida esta como el subconjunto de elementos de datos de la plantilla institucional que representa los campos que van a ser documentados y que representan el dato espacial, el producto o la información geográfica.

Por lo general, el trabajo de creación de las plantillas de metadatos se realiza una vez identificados los productos que se van a documentar y se produce con la intervención, generalmente, de un técnico, experto en el procedimiento de consecución del dato espacial, y otro, en la implementación del estándar y con la base de los parámetros que ofrece la "especificación técnica"<sup>10</sup> que se desarrolló para la elaboración del dato. La suma de estos componentes permite entregar una plantilla con información útil. Se presenta comúnmente la ausencia de los últimos dos factores que intervienen, pero también existen soluciones a estos problemas:

- En el caso de la falta de expertos en el tema de estándares de información geográfica, se cuenta con el apoyo del equipo profesional de la ICDE,

9 Aunque los perfiles de metadatos por lo general se realizan en ámbitos nacionales a partir de adopciones basados en la norma ISO 19115:2003 - Geographic information - Metadata, se justifica la creación de perfiles de estas normas, por ejemplo, para Colombia el perfil es la norma NTC 4611 segunda actualización, la cual para denominarse como un perfil de ISO debe cumplir con algunas pruebas para determinar su conformidad, estas las define la norma ISO 19106:2004 - Geographic information - Profiles. Sin embargo, para la generación de perfiles institucionales también se puede seguir el mismo procedimiento que se realiza para la generación de perfiles nacionales, en este caso el perfil será validado contra la norma de referencia, es decir, la NTC 4611 y se le aplicará las mismas pruebas de conformidad con lo que se le aplicó al perfil nacional.

10 Una especificación técnica es un documento con la descripción detallada de un producto o conjunto de datos que permiten crearlos, proveerlos y usarlos estableciendo los requerimientos y expectativas que se tienen por parte del cliente o usuario. NTC 5661 (2010).

para capacitar a las personas encargadas de los procesos. Alternativas como participar de los cursos virtuales y presenciales orientados a estándares de información geográfica y a las infraestructuras de datos espaciales es una gran posibilidad para solucionar el problema.

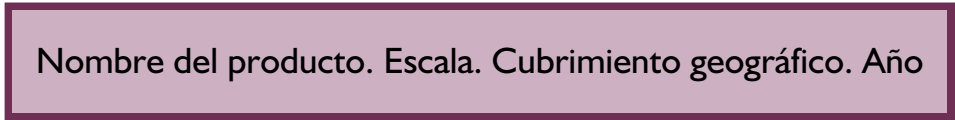
- La ausencia de especificaciones técnicas es un problema que para su solución requiere el apoyo de los grupos de trabajo de gestión de proyectos, por lo general, siempre existen documentos que determinan los parámetros y requerimientos de los usuarios y/o productores, por ejemplo, los convenios o contratos, las directrices y recomendaciones de una entidad, los marcos legales, las guías de implementación, manuales de procedimiento, instructivos, etc. Estos

documentos y su recopilación serán base para la generación de las plantillas de metadatos.

### 1.1 Normalización de los textos y citas en el metadato.

La plantilla de metadato debe contener los textos que intervendrán en la generación. Se normalizan los campos cuyo tipo de dato son texto libre, de tal forma que la estructura sea clara y permanezca inalterable durante el proceso de documentación. Por ejemplo, uno de los campos más importantes en el metadato es el campo Título, el cual permite identificar el producto por su nombre. Ha ocurrido que el solo nombre del producto no es suficiente para lograr identificar el dato, y se acostumbra asociarle a una estructura, como lo muestra la Figura 1.

**Figura 1.**  
Estructura del título del metadato  
Fuente: propia



Pero no basta con ofrecer en la plantilla la estructura del título, ya que se puede presentar que existan diferencias en la manera de presentar el mismo. Por ejemplo, en la Figura 2 y Figura 3<sup>11</sup>.

Como se puede apreciar en las Figura 2 y 3, los títulos hacen referencia a un mismo producto y cumplen con la estructura previamente definida, pero los textos tienen diferencias de forma

**Figura 2.**  
Reporte de metadato del mapa cartográfico básico de Melua - Meta  
Fuente: IGAC

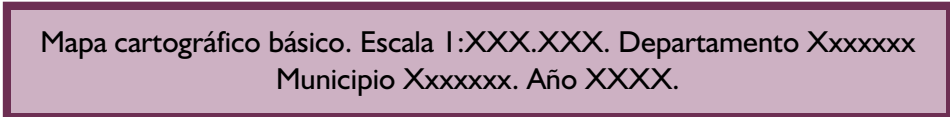
11 Las imágenes son reportes HTML de un metadato geográfico extraídas del Sistema Web de Administración de Metadatos SWAMI 3.0, catálogo de metadatos que soporta la ICDE.



**Figura 3.**  
Reporte de metadato del mapa cartográfico básico de Melua - Meta  
Fuente: IGAC

que pueden afectar la consulta que se realice hacia ellos en un futuro. Para la Figura 3, una persona no podría determinar a qué entidad territorial pertenece la plancha, en nuestro caso, es fácil identificar que Meta se refiere al departamento, pero para alguien ajeno al tema o que desconozca la división político-administrativa le resultaría difi-

cil. Igual sucede para la palabra Melua: si realizamos cualquier consulta en internet por este topónimo se encontrará: un planchón, un río, una vereda, etc. Por esto, es conveniente determinar los títulos y cualquier texto libre con más especificidad. Para el ejemplo anterior se soluciona el problema si se establece el título como se presenta en la Figura 4.



**Figura 4.**  
Ejemplo de título normalizado  
Fuente: propia

De esta forma se soluciona el inconveniente. Otro punto importante es la normalización de las citas para las fuentes de información, para el cual se debe realizar un trabajo similar al realizado con los campos de texto libre, determinando si se realiza la documentación de todos los insumos que intervienen o si se agrupan estos por su similitud. Por ejemplo, para el mismo mapa cartográfico, lo más probable es que se hayan utilizado como insumos fotografías aéreas pertenecientes a una faja de un vuelo, que cubra la zona geográfica en un cien mil. Pueden utilizarse más de 20 fotografías por cada vuelo, lo cual resultaría inmanejable si se genera una citación para cada foto, lo mejor es agruparlas y concebir un título genérico que agrupe a todas. Así se reduce el error que se pueda presentar en los textos de los campos.

### 1.2 Usando el Sistema Web de Administración de Metadatos SWAMI 3.0

Uno de los beneficios que ofrece el catálogo de metadatos SWAMI 3.0 es su capacidad de generar metadatos a partir de plantillas. SWAMI 3.0 las duplica y asigna un código al resultado, que hace referencia a un metadato que puede ser editado a partir de la interface de edición. El catálogo ofrece muchas ventajas, pero el interés se centra en la validación de la información cargada en el metadato.

A partir del cargue de la plantilla de metadato queda el registro XML<sup>12</sup> en la base de datos de la aplicación. Por medio de esta y del XSD<sup>13</sup> utilizado en la implementación del estándar NTC 4611 se realiza la validación en estructura del me-

12 Extensible Markup Language (XML). Este "metalenguaje" define la manera correcta para documentar la información con el fin de lograr compartir la información del metadato de manera eficaz y fiable, con otros programas de captura o editores.

13 XML Schema Definition (XSD). El XSD define la estructura que deben tener los campos y las restricciones de los mismos, por ejemplo, si se tiene que las fechas deben tener un formato predefinido de la manera AAAAMDD, el XSD permite estructurar el campo para que sea válido.

tadato asociado a la plantilla, es decir, la aplicación verifica que los elementos de dato estén cargados de acuerdo con los tipos de datos a los cuales deben estar asociados y que los elementos definidos

como obligatorios estén diligenciados, generando un reporte de validación para identificar los errores que tiene el metadato generado, como se puede apreciar en la Figura 5.

**Figura 5.**  
Reporte de validación de un metadato en SWAMI 3.0  
Fuente: IGAC

Tipo	Error	Nombre del Elemento	Ruta del Elemento
✘	El valor ingresado no es válido.	Oeste	Identificación - Identificación del Dato - Cubrimiento - Cubrimiento Geográfico - Identificación del Dato - Cubrimiento Geográfico - Coordenadas geográficas (límites - Oeste)
✘	El valor del elemento no es válido.	Oeste	Identificación - Identificación del Dato - Cubrimiento Geográfico - Coordenadas geográficas (límites - Oeste)
✘	El valor ingresado no es válido.	Este	Identificación - Identificación del Dato - Cubrimiento - Cubrimiento Geográfico - Coordenadas geográficas (límites - Este)
✘	El valor del elemento no es válido.	Este	Identificación - Identificación del Dato - Cubrimiento Geográfico - Coordenadas geográficas (límites - Este)
✘	El valor ingresado no es válido.	Sur	Identificación - Identificación del Dato - Cubrimiento - Cubrimiento Geográfico - Coordenadas geográficas (límites - Sur)
✘	El valor del elemento no es válido.	Sur	Identificación - Identificación del Dato - Cubrimiento Geográfico - Coordenadas geográficas (límites - Sur)

La opción de validación en el catálogo de metadato permite solo identificar los errores asociados a problemas en la estructura de los datos, pero cuando se tiene errores ortográficos, temáticos o conceptuales, técnicos, etc., es imposible identificarlos por medio de este procedimiento. Para eso es necesario realizar un control de calidad sobre la información misma, utilizando técnicas de muestreo.

## 2. Definiendo un procedimiento para evaluar la calidad de los metadatos

Para establecer un procedimiento válido para evaluar la calidad de los metadatos geográficos en un proceso de documentación se debe conocer en principio tres factores fundamentales:

Los criterios técnicos, los procesos y la normalización de los textos usados. Esta será la base para implementar un plan de calidad, que tendrá como objeto deter-

minar, según Capuz. *et al.*, 2000, los objetivos de la calidad, la responsabilidad en las diferentes fases del proyecto, los métodos, procedimientos e instrucciones de trabajo que deben aplicarse, los programas de inspección y la metodología para los cambios y modificaciones en el propio plan de calidad.

Así mismo, el procedimiento de evaluación se puede formular como guía metodológica que servirá de base temática en la ejecución de cualquier operativo<sup>14</sup> de migración. Si el plan de calidad define una secuencia lógica de pasos que enmarca las condiciones y las reglas de juego dentro del proceso, es indispensable describir específicamente las actividades a realizar por los grupos a cargo del control de calidad, para establecer los factores que influyen en la ejecución del mismo. Lo anterior permite proyectar las soluciones a fallas que puedan surgir en el desarrollo del proyecto. Estas han sido identificadas gracias a la implementación del plan de calidad y se expondrán en las conclusiones del presente artículo.

14 Un operativo de migración es el conjunto de actividades técnicas realizadas en una base de datos de metadatos, para actualizarlos a una versión diferente del estándar que los soporta.

## 2.1 Metodología

### *Determinación de los lotes*

Los lotes hacen referencia a la cantidad de metadatos generados o verificados dentro de un operativo de migración. Existen innumerables formas de identificar el lote y la forma de clasificarlos dependerá de la necesidad que tenga el técnico en la evaluación de calidad. En el caso de los operativos de captura de metadatos se ha logrado establecer que los mejores criterios para determinar un lote son:

- Cantidad de metadatos generados por técnico

Permite establecer en el muestreo simple un tamaño de lote asociado a una muestra.

- Fecha de captura

Permite identificar los lotes diarios, teniendo en cuenta que estos procesos obedecen a posibles errores inducidos por factores endógenos humanos, es

decir, casos asociados a la emotividad o al estado físico o al ambiente laboral.

- Responsable de captura

Permite identificar por técnico los lotes asociados a la producción.

- Por tipo de producto

Para agrupar los conjuntos de datos, de acuerdo con sus características temáticas, facilitando la revisión de los elementos de dato.

## 2.2 Determinación del Nivel de Inspección

En las pruebas realizadas se concluyó que un nivel general de inspección tipo II<sup>13</sup> es el que mejor se comporta, ya que permite realizar los cambios necesarios dentro del proceso cuando las evaluaciones de la muestra sugieren hacer cambios en el nivel de inspección. En el desarrollo del control de calidad se presentan cambios de dirección, según el siguiente esquema.

*Inspección Normal* ⇔ *Inspección Estricta* ⇔ *Inspección Uno a Uno*

Estos cambios obedecen a los lineamientos establecidos en los planes de muestreo, y se ejecutan cuando se encuentran errores sustanciales en

los lotes revisados, y se devuelve el procedimiento de acuerdo con la disminución del error encontrado de la siguiente forma:

*Inspección Uno a Uno* ⇔ *Inspección Estricta*

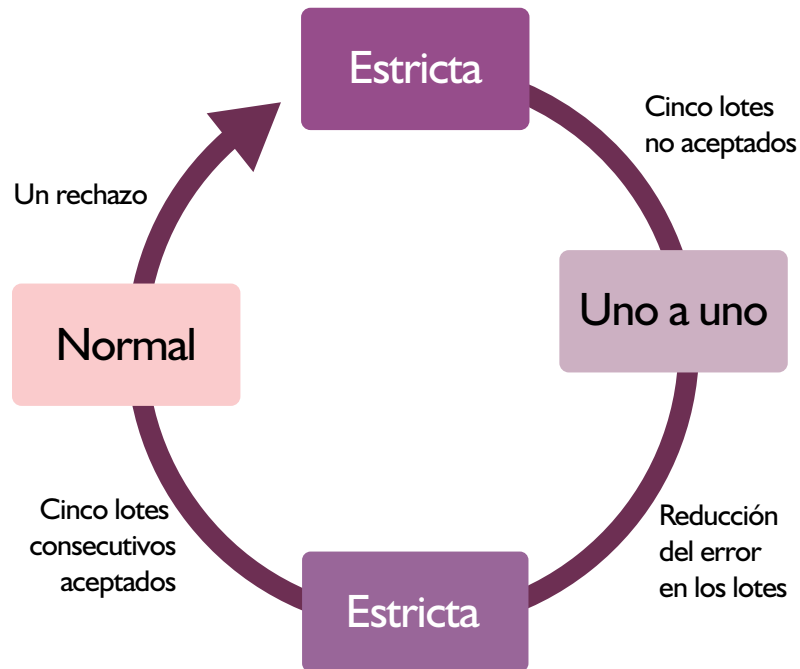
La inspección uno a uno se realiza en todos los metadatos del lote. Si no se encuentran errores adicionales se pasa a una inspección estricta. Por último, el ciclo concluye con el cambio.

En la cual se evidencia porque no se han encontrado errores en los lotes aceptados.

El procedimiento se resume en la Figura 6.

13 Un nivel de inspección de este concepto hace referencia a la relación entre el tamaño del lote y el tamaño de la muestra. Existen tres tipos de niveles de inspección, siendo el I el más riguroso y III el más general.

## *Inspección Estricta* ⇔ *Inspección Normal*



**Figura 6.**  
Ejemplo de  
procedimiento de  
inspección  
Fuente: propia

En la Figura 6 se muestra un ejemplo de procedimiento de inspección para un nivel de inspección tipo II; nótese que de la revisión normal se pasa a una estricta por encontrarse un error en la muestra utilizada. Al cambiar a la revisión estricta se revisarán más datos, es decir que el tamaño de la muestra aumentará; si no se encuentran errores en la nueva muestra se puede regresar a la revisión normal. Si se encuentran fallas en cinco lotes, se debe pasar a una revisión Uno a Uno y el universo se convierte en la muestra. Cuando se note una reducción de los errores presentes se puede bajar a una revisión estricta. No se puede pasar de una revisión Uno a Uno a una revisión Normal, se debe pasar antes por una revisión Estricta.

### 2.3 Muestreo simple

El muestreo a utilizar se basó en la Norma Técnica Colombiana NTC-ISO 2859 -1<sup>16</sup>, determinando usar la misma cantidad

de ítems que los utilizados en la muestra especificada en el plan.

Un plan de muestreo simple se define mediante tres elementos: el tamaño de la muestra, el número de aceptación y el número de rechazo. El plan se lleva a cabo, tomando del lote aleatoriamente el número de unidades requeridas de producto que conforma el tamaño de la muestra. Estas unidades de producto así escogidas se denominan "unidades de muestra" y en conjunto se conocen como "muestra".

La muestra se inspecciona y se cuenta el número de unidades defectuosas descubiertas, si dicho número es inferior o igual al número de aceptación se aprueba todo el lote, solamente se rechazarán los elementos de la muestra que se encontraron defectuosos. Si por otra parte, el número de unidades defectuosas es igual o superior al número de rechazo, se rechaza todo el lote.

16 La NTC presenta un sistema de muestreo para aceptación de muestras por inspección de atributos, y están orientados a aplicarse a ítems terminados, componentes y materias primas, operaciones, materiales en proceso, suministros en existencia, operaciones de mantenimiento, datos o archivos, procedimientos administrativos.

## 2.4 Determinando el tamaño de la muestra

En las tablas de la norma NTC-ISO 2859-1 se relaciona el tamaño de la muestra según la cantidad de productos del lote a inspeccionar. Un lote es la cantidad de metadatos revisados, generados o actualizados, basados en uno de los cuatro criterios definidos (cantidad de metadatos generados por técnico, fecha de captura, responsable de captura o tipo de producto). Para seleccionar la muestra, es decir, los metadatos que se tomarán para evaluar, se aplicará un método alea-

torio por medio de la fórmula en Excel. (=aleatorio.entre(1;100)).

Se adjudica un consecutivo a cada metadato del lote, el número resultante de aplicar la fórmula será el escogido como parte de la muestra. El tamaño de la muestra como se expuso anteriormente lo define la tabla. Por ejemplo, para un lote de 16 metadatos que obedece al rendimiento diario de un técnico de captura la muestra equivalente en un nivel general de inspección tipo II está determinado por la letra C, como se aprecia en la Figura 7.

TAMAÑO LOTE		Niveles de inspección especiales				Niveles generales de inspección		
		S1	S2	S3	S4	I	II	III
2	8	A	A	A	A	A	A	B
9	15	A	A	A	A	A	B	C
16	25	A	A	B	B	B	C	D
26	50	A	B	B	C	C	D	E
51	90	B	B	C	C	C	E	F
91	150	B	B	C	D	D	F	G
151	280	B	C	D	E	E	G	H
281	500	B	C	D	E	F	H	J
501	1200	C	C	E	F	G	J	K
1201	3200	C	D	E	G	H	K	L
3201	10000	C	D	F	G	J	L	M
10001	35000	C	D	F	H	K	M	N
35001	150000	D	E	G	J	L	N	P
150001	500000	D	E	G	J	M	P	Q
más de	500001	D	E	E	J	N	Q	R

Figura 7. Tabla 1. Letra código de tamaño de la muestra NTC-ISO 2859-1 Fuente: NTC-ISO 2859-1

Letra código tamaño de la muestra	Tamaño de la muestra	Nivel de calidad aceptable (NCA), en porcentaje de elementos no conformes y no conformidades por 100 unidades (inspección normal)																									
		0,010	0,015	0,025	0,040	0,065	0,10	0,15	0,25	0,40	0,65	1,0	1,5	2,5	4,0	6,5	10	15	25	40	65	100	150	250	400	650	1000
		Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re
A	2																										
B	3																										
C	5																										
D	8																										
E	13																										
F	20																										
G	32																										
H	50																										
J	80																										
K	125																										
L	200																										
M	315																										
N	600																										
P	800																										
Q	1250																										
R	2000																										

▼ = Utilizar el primer plan de muestreo bajo la flecha. Si el tamaño de la muestra es igual o excede el tamaño del lote, efectuar el 100% de la inspección  
 ▲ = Utilizar el primer plan de muestreo por encima de la flecha.  
 Ac = Valor de aceptación  
 Re = Valor de rechazo

Figura 8. Tabla 2.A. Planes de muestreo simple para inspección normal NTC-ISO 2859-1 Fuente: NTC-ISO 2859-1

En la Figura 8 se puede apreciar que la norma define para la letra C un tamaño de muestra de 5, en este caso los asociamos a 5 metadatos. Con un respectivo nivel de aceptación de 1 error y de rechazo de 2 errores. Lo anterior se interpreta de la siguiente manera, en el caso que en la muestra se identifique solo un error este se corrige y se acepta el lote, pero en el caso de que se encuentren dos errores este se rechaza y se pasa a una inspección estricta.

## 2.5 Determinación de las variables a verificar

Previo a la migración de los metadatos se identifican los campos que van a ser objeto de edición; para esto es necesario realizar un estudio del estado de la información, cuyo resultado debe establecer la solución a los posibles casos que se presenten. Por ejemplo, se puede presentar nombres de dátums diferentes; esto se debe normalizar para ofrecer un único nombre que estará asociado a un código EPSG<sup>17</sup>.

Las variables son definidas a partir de aquellos textos que cambian o que son editados en el proceso de documentación, todo campo al que se ingresa para cambiar el dato debe ser objeto de revisión, de aquí la importancia de contar con plantillas de metadato que permitan identificar previamente aquellos elementos de dato que se actualizan. La eficacia de las plantillas se revela en esta fase ya que se puede desconocer aquellos campos plantillados a los cuales no se ingresa y están exentos de errores de digitación o temáticos.

## Conclusiones

Un proceso efectivo de calidad en un proyecto de generación de metadatos debe estar sustentado por parámetros estadísticos que permitan identificar los errores presentes en una muestra para asociarlos al universo de metadatos documentados. Esto reduce en gran medida los costos del proyecto y da confianza al productor de los datos.

El uso de plantillas y perfiles institucionales de metadato reduce los errores que se presentan en la documentación de los metadatos, pues restringe el acceso a los elementos de dato que no se van a utilizar por haberse definido los textos con anterioridad.

Se deben implementar en las evaluaciones de calidad procedimientos automatizados, la facilidad que SWAMI ofrece a sus usuarios en cuanto a validación de metadatos en estructura debe ser explotada y difundida. Con esto se contribuye a que la información dispuesta en el catálogo cumpla con un estándar nacional y garantice el intercambio de metadatos con otras IDE.

Aparte del beneficio de tener metadatos con calidad, un procedimiento técnico contribuye a la normalización del proceso de migración, fundamentando todas las fases de evaluación de calidad bajo los parámetros que suministran las NTC y solucionando problemas tan sencillos como obtener los tamaños de muestras y la determinación de los lotes, que afectan el desarrollo de las actividades posteriores. Tener en cuenta desde un principio estos elementos previene vacíos dentro del procedimiento y da certeza sobre las acciones a realizar cuando se presenta un comportamiento inesperado.

15 La codificación de la European Petroleum Survey Group (EPSG) es la utilizada en la documentación de los sistemas de referencia en SWAMI, la cual reduce toda la información asociada a un código.

## Referencias bibliográficas

- CEBRÍAN, De Miguel J. Antonio (1992). Información Geográfica y Sistemas de Información Geográfica. Servicio de publicaciones de Cantabria, pág. 25.
- GRANELL, C. Gould. M. (2006). Avances en la Infraestructura de Datos Espaciales. Cantabria: Publicacions de la Universitat Jaume I, pág. 8.
- CAPUZ, S., GÓMEZ, E., TORREALBA, A., FERRER, P., GÓMEZ, T., VIVANCO, J. (2000). Cuadernos de ingeniería de proyectos III: Dirección, Gestión y Organización de Proyectos. Valencia: Universidad Politécnica de Valencia, pág. 56.
- INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS Y CERTIFICACIÓN (ICONTEC). Información geográfica. Metadato Geográfico. NTC 4611 segunda actualización. Bogotá, D. C.: ICONTEC, 2011.
- . Información geográfica. Procedimiento y medidas para la evaluación de calidad. NTC 5660. Bogotá, D. C.: ICONTEC, 2010.
- . Información geográfica. Principios básicos de calidad. NTC 5043, primera actualización. Bogotá, D. C.: ICONTEC, 2010.
- . Procedimientos de muestreo para la inspección de atributos. Parte 1: Planes de muestreo determinados por el nivel aceptable de calidad (NAC) para inspección lote a lote. NTC 2859-1. Bogotá D.C.: ICONTEC, 2002.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Geographic information - Quality Principles. ISO 19113:2002. Ginebra: ISO, 2002.
- . Geographic information - Terminology. ISO 19106:2003. Ginebra: ISO, 2003.
- . Geographic information - Metadata. ISO 19115:2003. Ginebra: ISO, 2003.