

**INFORMACIÓN GEOGRÁFICA. CALIDAD DE LOS DATOS. PARTE 1: REQUISITOS
GENERALES**

**Correspondencia: ISO 19157-1:2023
I.C.S.: 35.240.70**

**DIRECCIÓN GENERAL DE NORMAS Y TECNOLOGÍA INDUSTRIAL (DGNTI)
ORGANISMO NACIONAL DE NORMALIZACIÓN DE LA REPÚBLICA DE PANAMÁ.
(Título II ley 23 de 15 de julio de 1997)**

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Contenido

1. OBJETO Y CAMPO DE APLICACIÓN	6
2. NORMAS PARA CONSULTA	6
3. TÉRMINOS Y DEFINICIONES	7
4. ABREVIATURAS DE TÉRMINOS Y DE PAQUETES	10
5. CONFORMIDAD	11
6. REQUISITOS GENERALES PARA LA CALIDAD DE LA INFORMACIÓN GEOGRÁFICA	12
7. VISIÓN GENERAL DE LA CALIDAD DE LOS DATOS	16
8. COMPONENTES DE LA CALIDAD DE LOS DATOS	18
9. MEDIDAS DE CALIDAD DE LOS DATOS	32
10. EVALUACIÓN DE LA CALIDAD DE LOS DATOS	38
11. INFORME DE LA CALIDAD DE LOS DATOS	43
12. REQUISITO PARA LA CODIFICACIÓN XML	48
Anexo A.....	49
ANEXO B (Normativo) Conceptos sobre la calidad de los datos y su uso.....	51
ANEXO C (Normativo) Diccionario de datos para la calidad de los datos	61
ANEXO D (Informativo) Evaluación e informe de la calidad de los datos	79
ANEXO E (informativo) Métodos de muestreo para la evaluación de la calidad de los datos	108
ANEXO F (informativo) Métodos de muestreo para la evaluación de la calidad de los datos	119
ANEXO G (informativo) Agregación de resultados de la calidad	131
ANEXO H (Normativo) Descripción de la codificación XML	134
Anexo I	135

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

PREFACIO

La Dirección General de Normas y Tecnología Industrial (DGNTI), del Ministerio de Comercio e Industrias (MICI) es el Organismo Nacional de Normalización encargado por el estado del proceso de Normalización Técnica, Evaluación de la Conformidad y Certificación de Calidad.

Esta Norma Técnica en su etapa de proyecto, ha sido sometida a un período de discusión pública de treinta (30) días.

La Norma Técnica DGNTI ISO 19157-1:2023 ha sido oficializada por el Ministerio de Comercio e Industrias mediante la resolución N° XXX del XXXXX de XXXX de XXXX, y publicada en Gaceta Oficial N° XXXXX del XX de XXXXX de XXXX.

Esta Norma DGNTI ISO 19157-1:2023 es equivalente a la Norma ISO 19157-1:2023.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Prólogo de la versión en español

Esta Norma Internacional es una adopción directa de la Norma UNE EN ISO 19157-1:2024, que a su vez adopta la Norma Internacional ISO 19157-1 y ha sido traducida por Comité Europeo de Normalización, en el que participan representantes de los Organismos Nacionales de Normalización y representantes del sector empresarial de los siguientes países:

Alemania, Austria, Bélgica, Bulgaria, Chipre, Croacia, Dinamarca, Eslovaquia, Eslovenia, España, Estonia, Finlandia, Francia, Grecia, Hungría, Irlanda, Islandia, Italia, Letonia, Lituania, Luxemburgo, Malta, Noruega, Países Bajos, Polonia, Portugal, Reino Unido, República Checa, República de Macedonia del Norte, Rumania, Serbia, Suiza y Turquía.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

INTRODUCCIÓN

Los datos geográficos se comparten, intercambian y utilizan cada vez más para propósitos diferentes de los inicialmente pretendidos por sus productores. Por ello es vital tener información acerca de la calidad de los datos geográficos disponibles, para orientar los procesos de selección en los que el valor de los datos está directamente relacionado con su calidad. Un usuario de datos geográficos puede tener múltiples conjuntos de datos entre los que elegir. De este modo, es necesario comparar la calidad de los conjuntos de datos para determinar cuál de ellos es el que satisface mejor los requerimientos del usuario.

El propósito de describir la calidad de los datos geográficos es facilitar la comparación y selección más adecuada para las necesidades o requisitos de una aplicación. Las descripciones completas de la calidad fomentarán que se compartan, intercambien y utilicen los conjuntos de datos más apropiados. La información acerca de la calidad de datos geográficos permite a sus productores evaluar la adecuación de un conjunto de datos a los criterios establecidos en las especificaciones de producto y ayuda a los usuarios a determinar la capacidad de dicho producto para satisfacer los requisitos de su aplicación particular. Para el propósito de esta evaluación se utilizan de manera consistente procedimientos claramente definidos.

Para facilitar las comparaciones es esencial que los resultados de calidad se expresen de forma comparable y que exista un entendimiento común de las medidas de calidad de datos que se han utilizado. Estas medidas de calidad de los datos proporcionan descriptores de la calidad de los datos geográficos mediante la comparación con el universo de discurso. El uso de medidas incompatibles hace que no sean posibles las comparaciones sobre la calidad de datos. Este documento normaliza las componentes y estructuras de las medidas de calidad de datos y define las que más suelen utilizarse.

Este documento considera que productores y usuarios pueden potencialmente ver la calidad de datos desde diferentes perspectivas. Los niveles de conformidad de la calidad de datos pueden establecerse usando bien las especificaciones de producto realizadas por el productor o bien los requerimientos del usuario sobre la calidad de los datos. Si el usuario requiere más información sobre la calidad que aquella suministrada por el productor, aquél puede seguir el flujo del proceso de evaluación de la calidad de datos realizado por este para obtener información adicional. En ese caso, los requerimientos de usuario se consideran como especificaciones de producto con el propósito de usar el flujo del proceso del productor.

El objetivo de este documento es proveer un marco para la definición de la calidad de los datos geográficos. Esto incluye los principios para la evaluación de la calidad, un modelo conceptual para el tratamiento de la información sobre la calidad, una estructura y contenido de las medidas de calidad de datos, y pautas para informar sobre una evaluación de la calidad. El marco ampliable, con reglas sobre cómo añadir medidas de calidad de datos. También prevé las dimensiones complejas de la calidad de los datos.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

INFORMACIÓN GEOGRÁFICA. CALIDAD DE LOS DATOS. PARTE 1: REQUISITOS GENERALES

1. OBJETO Y CAMPO DE APLICACIÓN

Este documento establece los principios para describir la calidad de datos geográficos. Esta norma:

- Define un sistema sopesado de componentes para describir la calidad de datos;
- Define el proceso para definir componentes adicionales de campos específicos para describir la calidad de los datos;
- Especifica las componentes y la estructura del contenido de las medidas de calidad de los datos;
- Describe los procedimientos generales para evaluar la calidad de los datos geográficos:
- Establece principios para informar sobre la calidad de los datos.

Este documento es aplicable por productores de datos que proporcionan información referente a la calidad para describir y valorar el grado de adecuación con que un conjunto de datos se ajusta a las especificaciones de productos. También es aplicable por los usuarios que tratan de determinar si unos datos concretos son de suficiente calidad para su aplicación particular.

Este documento no pretende definir niveles mínimos de aceptación para la calidad de los datos geográficos. Esta información suele presentarse como un requisito en una especificación de producto de datos, definida, por ejemplo, de acuerdo con la Norma ISO 19131.

2. NORMAS PARA CONSULTA

En el texto se hace referencia a los siguientes documentos de manera que parte o la totalidad de su contenido constituyen requisitos de este documento. Para las referencias con fecha, solo se aplica la edición citada- Para los referencias sin fecha se aplica la última edición (incluida cualquier modificación de esta).

ISO 19103:20215, *Información geográfica. Lenguaje de esquema conceptual.*

ISO 19109:2015, *Información geográfica. Reglas para esquemas de aplicación.*

ISO 19115-1, *Información geográfica. Metadatos. Parte 1: Fundamentos*

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

3. TÉRMINOS Y DEFINICIONES

Para los fines de este documento, se aplican los términos y definiciones siguientes.

ISO e IEC mantienen bases de datos terminológicas para su utilización en normalización en las siguientes direcciones:

- Plataforma de búsqueda en línea de ISO: disponible en <https://www.iso.org/obp>
- Electropedia de IEC: disponible en <http://www.electropedia.org/>

3.1. Exactitud

Grado de concordancia entre el resultado de un ensayo o una medición y el valor verdadero.

Nota 1: En este documento, el valor verdadero puede ser un valor de referencia que es aceptado como verdadero.

[FUENTE: ISO 3534-2:2006, 3.3.1, modificado. Las notas 1, 2 y 3 se han eliminado. Se ha añadido una nueva nota.]

3.2 Conformidad

Cumplimiento de un requisito.

Nota 1: Cuando no haya ambigüedad se puede omitir “conformidad. Por ejemplo, “informe de prueba” es lo mismo que “informe de prueba de conformidad”.

3.3 Nivel de conformidad de la calidad

Valor umbral o conjunto de valores umbral, para los resultados de la calidad de los datos, empleados para determinar lo bien que un conjunto de datos se adecua a los criterios expuestos en sus especificaciones de producto de datos o en los requerimiento de usuario.

3.4 Corrección

Correspondencia con el universo de discurso.

3.5 Cobertura

Fenómeno que actúa como una función que devuelve valores de su rango para cualquier posición directa dentro de su dominio.

3.6 Especificaciones de producto de datos

Especificación de un producto de datos junto con la información adicional que permite su creación, suministro y utilización por otras partes.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

3.7 Calidad de los datos

Grado con el que un conjunto de características inherentes de los datos cumple unos requisitos.

[FUENTE: ISO 8000-2:2022, 3.8.1, modificado. La nota 1 se ha eliminado].

3.8 Medida de la calidad del dato

Variable a la que se asigna un valor como resultado de la medición de una característica de calidad de datos.

[FUENTE: ISO/IEC 25012:2008, 4.5, modificado. La nota 1 se ha eliminado]

3.9 Unidad de calidad de los datos

Combinación de un ámbito y elementos de la calidad de los datos.

3.10 Conjunto de datos

Colección identificable de datos.

NOTA 1: Un conjunto de datos puede ser un agrupamiento más pequeño de datos el cual, aunque limitado por alguna restricción tal como la extensión espacial o el *tipo de objeto geográfico*, se puede localizar físicamente dentro de un conjunto de datos más grande. Teóricamente, un conjunto de datos puede ser tan pequeño como un único *objeto geográfico* o un atributo de *objeto geográfico* contenidos dentro de un conjunto de datos más grande. Un mapa impreso o una carta pueden considerarse también como conjuntos de datos.

3.11 Serie de conjuntos de datos:

Colección de conjuntos de datos que comparten características comunes.

3.12 Objeto geográfico

Abstracción de un fenómeno del mundo real.

NOTA 1: Un objeto geográfico puede darse como un tipo como una instancia. Se usará el tipo de objeto geográfico o instancia de objeto geográfico solo cuando se desea referirse a uno de ellos.

3.13 Atributo de objeto geográfico

Características de un objeto geográfico.

NOTA 1: Un atributo de objeto geográfico tiene un nombre, un tipo de datos y un dominio de valores asociado. Un atributo de objeto geográfico para una instancia de objeto geográfico también tiene un valor de atributo tomado del dominio de valores.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

3.14 Instancia de objeto geográfico

Individuo de un tipo de objeto geográfico dado que tiene valores específicos de los atributos de objeto geográfico.

[FUENTE: ISO 19101-1:2014, 4.1.14]

3.15 Operación de objeto geográfico:

Operación que cada instancia de un tipo de objeto geográfico puede realizar.

[FUENTE: ISO 19110:2016, 3.7, modificado. El ejemplo y la nota 1 se han eliminado]

3.16 Tipo de objeto geográfico

Clase de objetos geográficos que tienen características comunes.

[FUENTE: ISO 19156:2011, 4.7]

3.17 Datos geográficos

Datos que implícita o explícitamente se refieren a una localización relativa a la Tierra.

[FUENTE: ISO 19109:2015, 4.13, modificado. La nota 1 se ha eliminado.]

3.18 Ítem

Lo que se puede describir y considerar individualmente.

NOTA 1: Un ítem puede ser cualquier parte de un conjunto de datos, tal como un objeto geográfico, una relación entre objetos geográficos, un atributo de objeto geográfico, o una combinación de estos.

3.19 Linaje

Procedencia, fuente(s) y proceso(s) de producción utilizados en la producción de un recurso.

[FUENTE: ISO 19115-1:2014, 4.9]

3.20 Metadatos

Información acerca de un recurso.

[FUENTE: ISO 19115-1:2024, 4.10]

3.21 Metacalidad

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Información que describe la calidad de la calidad de los datos.

3.23 Evaluación de la calidad:

Examen sistemático del grado en que un objeto geográfico es capaz de cumplir requisitos especificados.

[FUENTE: ISO/IEC/IEEE 24765:2017, 3.3267, modificado. La nota 1 se ha eliminado]

3.24 Registro

Conjunto de ficheros que contienen identificadores asignados a ítems con descripciones de los ítems asociados.

3.25 Requisitos

Necesidad o expectativa establecida, generalmente implícita u obligatoria.

[FUENTE: ISO 9000:2015, 3.6.4, modificado. Las notas 1-6 se han eliminado]

3.26 Informe de evaluación de la calidad; informe de la calidad

Documento de texto libre que proporciona información completa y detallada sobre las evaluaciones, resultados y medidas de calidad de los datos utilizados.

3.27 Incertidumbre; Incertidumbre de medición

Parámetro, asociado con el resultado de una medición, que caracteriza la dispersión de los valores que podrían atribuirse razonablemente al mensurando.

NOTA 1: En general, la incertidumbre de medición incluye numerosas componentes. Algunas pueden evaluarse a partir de la distribución estadística de los valores que proceden de las series de mediciones y pueden caracterizarse por desviaciones típicas. Los otros componentes pueden caracterizarse también por desviaciones típicas, evaluadas a partir de funciones de densidad de probabilidad basadas en la experiencia u otra información.

3.28 Universo de discurso

Vista del mundo real o hipotético que incluye todo aquello que es de interés.

[FUENTE: ISO 19101-1:2014, 4.1.38]

4. ABREVIATURAS DE TÉRMINOS Y DE PAQUETES

4.1 Abreviaturas de términos

RACD Resultado agregado de la calidad de los datos

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

LCA	Límite de calidad de aceptación
GML	Lenguaje geográfico de marcado (Geographic Markup Language)
UML	Lenguaje unificado de modelado (Unified Modeling Language)
XML	Lenguaje extensible de marcado (Extensible Markup Language)

4.2 Abreviaturas de paquetes

Las abreviaturas se usan para indicar el paquete que contiene una clase. Estas abreviaturas preceden a los nombres de las clases, conectados con un “_”. En la tabla 1 se presenta un listado de esas abreviaturas. La norma internacional en la que se localizan las clases se indica entre paréntesis.

Tabla 1 – Abreviaturas de paquetes

CI	Mención (Citation) [ISO 19115-1:2014]
DS	Conjunto de datos (Dataset) [ISO 19115-1:2014]
GFM	Modelo General de Objetos Geográficos (General Feature Model) [ISO 19109:2005]
LI	Linaje (Lineage) [ISO 19115-1:2024]
MD	Metadatos (Metadata) [ISO 19115-1:2014]

5. CONFORMIDAD

5.1 Generalidades

En este documento se definen dos clases de conformidad (véase 5.2 y 5.3). Las pruebas relacionadas se proporcionan en el conjunto de pruebas abstractas presentado en el anexo A.

Los requisitos, recomendaciones y permisos están marcados explícitamente y se les ha asignado un identificador de requisito, recomendación o permiso.

El nombre y la información de contacto de la agencia de mantenimiento de este documento se pueden encontrar en www.iso.org/maintenance_agencies.

5.2 Contenido de un modelo de la calidad de los datos

La tabla 2 describe la clase de conformidad para el contenido de una definición de la calidad de los datos.

Tabla 2 – Clase de conformidad del contenido

Clase de conformidad	https://standards.iso211.org/19157/-1/1/conf/content
Tipo de objetivo de normalización	Instancia de una definición de calidad de los datos, con independencia de la codificación de los datos
Dependencia	https://standards.iso211.org/19157/-1 (Lenguaje de esquemas conceptuales) https://standards.iso211.org/19157/-2 (Reglas para esquemas de aplicación) https://standards.iso211.org/19157/-1/1/ (Metadatos – Parte 1: Fundamentos)
Clase de requisitos	https://standards.iso211.org/19157/-1/1/req/content (véase capítulo 6)
Pruebas	Todas las pruebas del capítulo A.1.

5.3 Codificación XML de un modelo de calidad de los datos

En la tabla 3 se describe la clase de conformidad para la representación XML de un modelo de la calidad de los datos.

Tabla 3 – Clase de conformidad para la codificación XML

Clase de conformidad	https://standards.iso211.org/19157/-1/1/conf/xml
Tipo de objetivo de normalización	Documento XML representando un modelo de la calidad de los datos.
Dependencia	https://standards.iso211.org/19157/-1/1/conf/content
Clase de requisitos	https://standards.iso211.org/19157/-1/1/req/xml (Véase capítulo 12)
Pruebas	Todas las pruebas del capítulo A.2.

6. REQUISITOS GENERALES PARA LA CALIDAD DE LA INFORMACIÓN GEOGRÁFICA

6.1 Generalidades

Los capítulos 8 a 11 describen las componentes de la calidad de los datos basándose en el modelo de Lenguaje Unificado de Modelado (UML, *Unified Modeling Language*) definido en este documento, que forma parte del modelo armonizado ISO/TC 211. Además, se incluyen descripciones, requisitos y recomendaciones adicionales. En el anexo C se proporciona un diccionario de elementos en modelo UML.

CONSULTA PÚBLICA

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

6.2 Calidad de los datos – requisitos, recomendaciones y permisos generales

En la tabla 4 se ofrece una lista de los requisitos definidos en este documento.

Tabla 4 – Lista de requisitos

Clase de requisito		https://standards-isotc211.org/19157/-1/1/req/content
Tipo de objetivo de normalización		Instancia de una definición de calidad de los datos, con independencia de la codificación de los datos
Dependencia		https://standards-isotc211.org/19157/-/1/ (Lenguaje de esquemas conceptuales)
Dependencia		https://standards-isotc211.org/19157/-/2/ (Reglas para esquemas de aplicación)
Dependencia		https://standards-isotc211.org/19157/-/1/1/ (Metadatos. Parte 1: Fundamentos)
Requisito 1		https://standards.isotc211.org/19157/-1/1/req/content/dataQuality
Requisito 2		https://standards.isotc211.org/19157/-1/1/req/content/additionalQualityElement
Requisito 3		https://standards.isotc211.org/19157/-1/1/req/content/qualityMeasure
Requisito 4		https://standards.isotc211.org/19157/-1/1/req/content/additionalQualityMeasure
Requisito 5		https://standards.isotc211.org/19157/-1/1/req/content/dataQualityMetadata
Requisito 6		https://standards.isotc211.org/19157/-1/1/req/content/qualityEvaluationReport
Requisito 7		https://standards.isotc211.org/19157/-1/1/req/content/aggregatedResult
Requisito 8		https://standards.isotc211.org/19157/-1/1/req/xml

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Además de la clase de requisito de la tabla 4, en la tabla 5 se define una clase de recomendaciones y en la tabla 6 se define una clase de permiso. Estas recomendaciones y permisos, cuando se ponen en práctica, pueden contribuir a mejorar el contenido de una definición de la calidad de los datos. No obstante, las recomendaciones y permisos son opcionales y no tienen repercusión en los resultados de la prueba de conformidad; por lo tanto, la clase de recomendación de la tabla 5 y la clase de permiso de la tabla 6 no están relacionadas con ninguna de las clases de conformidad definidas en el capítulo 4.

Tabla 5 – Lista de recomendaciones

Clase de recomendación	https://standards.iso211.org/19157/-1/1/req/optionalContent
Tipo de objetivo de normalización	Instancia de una definición de calidad de los datos, con independencia de la codificación de los datos
Dependencia	https://standards.iso211.org/19157/-1/1/req/content
Recomendación 1	https://standards.iso211.org/19157/-1/1/rec/optionalContent/dateTime
Recomendación 2	https://standards.iso211.org/19157/-1/1/rec/optionalContent/identifier
Recomendación 3	https://standards.iso211.org/19157/-1/1/rec/optionalContent/additionalQualityMeasure
Recomendación 4	https://standards.iso211.org/19157/-1/1/rec/optionalContent/description
Recomendación 5	https://standards.iso211.org/19157/-1/1/rec/optionalContent/qualityEvaluationProcess
Recomendación 6	https://standards.iso211.org/19157/-1/1/rec/optionalContent/evaluationMethodType
Recomendación 7	https://standards.iso211.org/19157/-1/1/rec/optionalContent/reportReference
Recomendación 8	https://standards.iso211.org/19157/-1/1/rec/optionalContent/aggregatedResult
Recomendación 9	https://standards.iso211.org/19157/-1/1/rec/optionalContent/sameQualityElement
Recomendación	https://standards.iso211.org/19157/-

10	1/1/rec/optionalContent/differentQualityElement
Recomendación 11	https://standards.iso211.org/19157/-1/1/rec/optionalContent/qualityEvaluationReport
Recomendación 12	https://standards.iso211.org/19157/-1/1/rec/optionalContent/derivedResultMetadata
Recomendación 13	https://standards.iso211.org/19157/-1/1/rec/optionalContent/hierarchy

En la tabla 6 se presenta una lista con los permisos definidos en este documento.

Tabla 6 – Lista de permisos

Clase de permiso	https://standards.iso211.org/19157/-1/1/req/permitedContent
Tipo de objetivo de normalización	Instancia de una definición de calidad de los datos, con independencia de la codificación de los datos
Dependencia	https://standards.iso211.org/19157/-1/1/req/content
Permiso 1	https://standards.iso211.org/19157/-1/1/per/permitedContent/additionalQualityElement
Permiso 2	https://standards.iso211.org/19157/-1/1/per/permitedContent/conformanceResult
Permiso 3	https://standards.iso211.org/19157/-1/1/per/permitedContent/sourceReference

7. VISIÓN GENERAL DE LA CALIDAD DE LOS DATOS

Trabajar con la calidad de los datos incluye los siguientes puntos:

- Comprender los conceptos de la calidad de los datos relacionados con los datos geográficos. El anexo B proporciona una descripción de los conceptos de la calidad de los datos utilizados para establecer los componentes que describen la calidad de los datos geográficos.
- Definir los niveles de conformidad de a calidad de los datos según las especificaciones de producto de datos o basadas en los requerimientos de usuario. El establecimiento de especificaciones de producto de datos se describe en la Norma ISO 19131.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Especificar los aspectos de la calidad en los esquemas de aplicación.
- Evaluar la calidad de los datos y la metacalidad.
- Informar sobre la calidad de los datos y la metacalidad.

NOTA 1: El desarrollo de esquemas de aplicación se describe en la Norma ISO 19109.

Una evaluación de la calidad de los datos puede aplicarse a una serie de conjuntos de datos, a un conjunto de datos, o a un subconjunto de datos dentro de este, que compartan características comunes, de forma que pueda evaluarse su calidad.

Los elementos de la calidad de los datos y sus descriptores describen el grado de adecuación de un conjunto de datos a los criterios establecidos en sus especificaciones de producto, o a los requerimientos de usuario, y proporcionan información cuantitativa sobre la calidad.

Requisito 1	https://standards.iso/211.org/19157-1/1/req/content/dataQuality
	La calidad de los datos debe describirse de conformidad con las componentes de la calidad de los datos y las medidas de calidad de los datos según lo definido por las figuras UML en los capítulos 8 y 9 de este documento.

Cuando la información sobre la calidad de los datos describe datos que se han creado sin especificaciones de producto detalladas, o con especificaciones que carecen de medidas cuantitativas y descriptores, la calidad de los datos puede evaluarse subjetivamente de forma no cuantitativa con un resultado descriptivo (véase 8.5.4.4) utilizado para informar sobre el resultado de la evaluación para cada uno de los elementos de la calidad de los datos.

El propósito, uso y linaje proporcionan cierta información relacionada con la calidad. Esa información se consigna como metadatos de acuerdo con la Norma ISO 19115-1.

NOTA 2: El 'propósito' describe la justificación para la creación del conjunto de datos y contiene información sobre su utilización prevista, que puede potencialmente no ser la misma que la real. El 'linaje' describe la historia de un conjunto de datos y refiere su ciclo de vida, desde su captura y adquisición, pasando por su compilación y derivación, hasta su forma actual. Esta información general, no cuantitativa, es ilustrativa para los usuarios y puede ayudar en la evaluación de la calidad de un conjunto de datos, especialmente en aquellos casos en los que se utilizan para una aplicación particular que difiere de la prevista (véase también 10.2.3).

Este documento reconoce que los elementos de la calidad e los datos pueden tener asociada una calidad denominada metacalidad. La metacalidad describe la calidad

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

de los resultados de la calidad de los datos según una características definidas.

NOTA 3: El concepto de metacalidad se describe en el apartado 8.3.7 y sus descriptores se definen en el apartado 8.5.5.

La figura 1 proporciona una visión general de la información sobre la calidad de los datos.

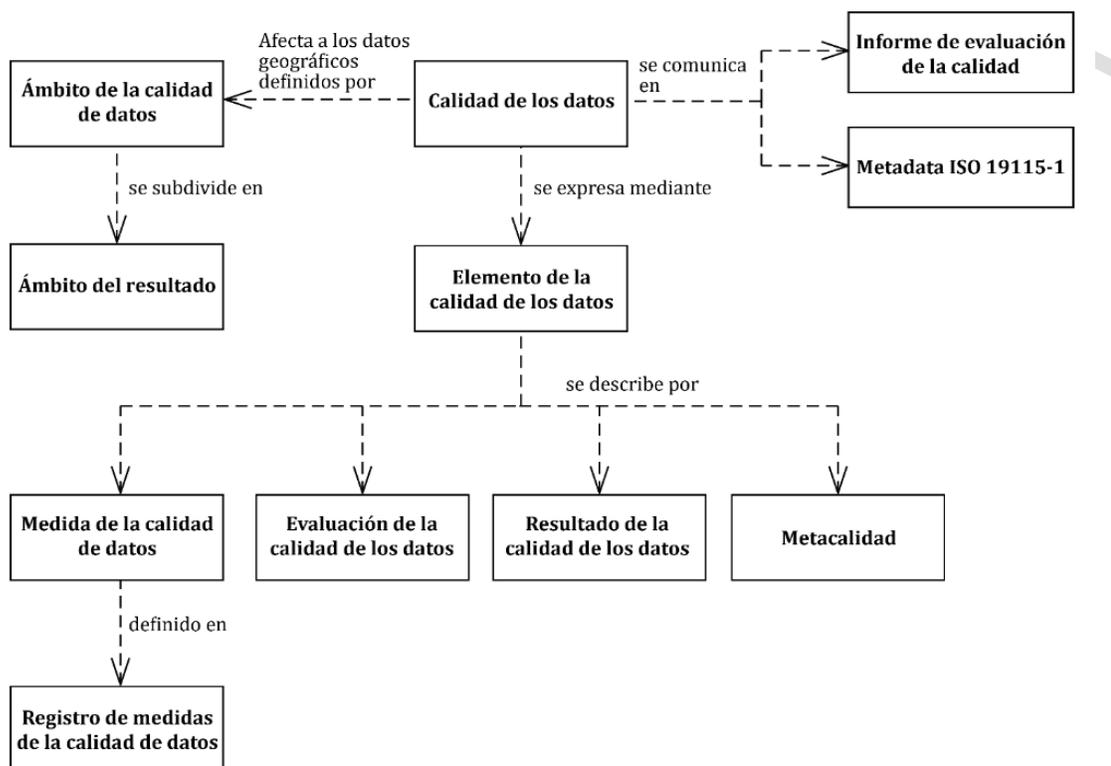


Figura 1 – Modelo conceptual de la calidad para los datos geográficos

8. COMPONENTES DE LA CALIDAD DE LOS DATOS

8.1 Visión General de las componentes

Las componentes de la calidad de datos se describen en el capítulo 8. La figura 2 presenta una visión general de las componentes y las conexiones entre ellas. Para más detalles sobre las componentes y sus atributos véase el diccionario de datos definido en el anexo C.

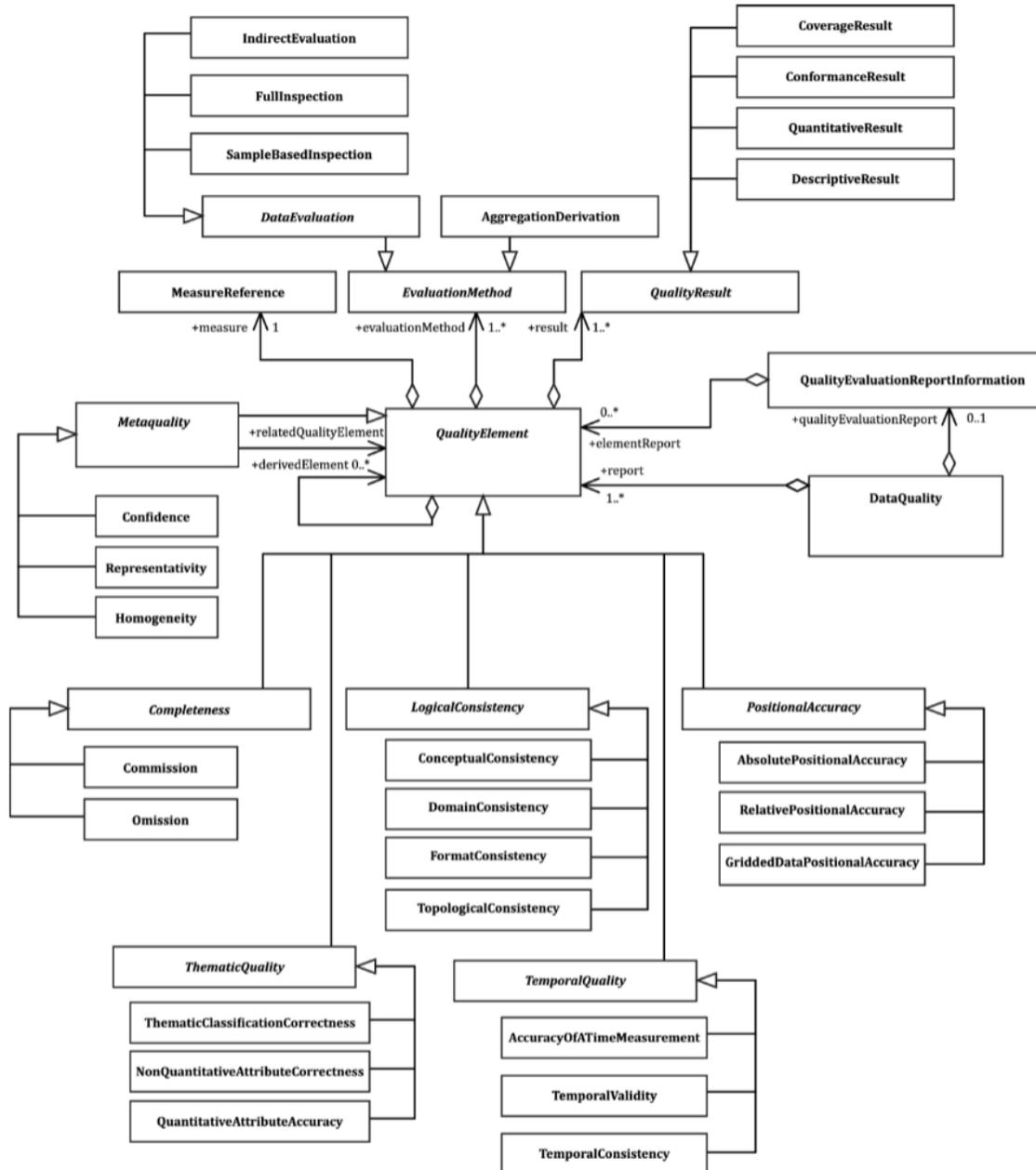


Figura 2 – Visión general de las componentes de la calidad de los datos

8.2. UNIDAD DE CALIDAD DE LOS DATOS

Al describir la calidad de los datos geográficos pueden considerarse diferentes elementos de calidad y diferentes subconjuntos de datos. Para describirlos se utilizan las unidades de la calidad de los datos. Una unidad de calidad de los datos es la combinación de un ámbito y uno o más elementos de la calidad de los datos; véase la figura 3 y la tabla C.1.

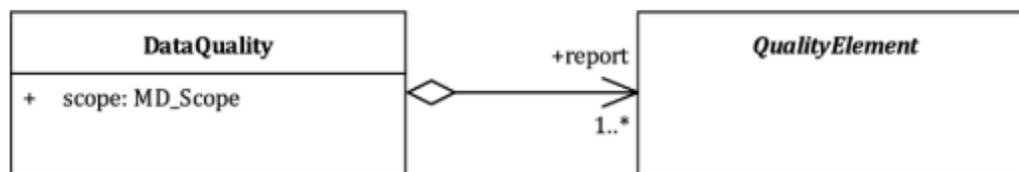


Figura 3 – Unidad de calidad de los datos

El ámbito de la(s) unidad(es) de calidad de los datos especifica la extensión, espacial o temporal, o la(s) características(s) común(es) que identifican los datos cuya calidad tiene que establecerse y evaluarse.

Puesto que los ámbitos varían frecuentemente para los diferentes elementos de la calidad de datos individuales, un informe sobre la calidad de los datos (como metadatos o como informe de evaluación de la calidad) puede abarcar varias unidades. Por ejemplo, tales ámbitos diferentes pueden estar especialmente separados, superpuestos o incluso compartir la misma extensión.

NOTA: Si un conjunto de datos es el resultado de la integración de conjuntos de datos de diferentes ámbitos, valores de la calidad de los datos o ponderaciones de elementos de la calidad, ese nuevo conjunto de datos necesita que se establezca su nueva unidad de calidad de los datos, es decir, definir su propio ámbito de calidad de los datos y especificar los elementos de la calidad que interesan.

Los siguientes son ejemplos de lo que define un ámbito de calidad de los datos (véase también MD_Scope en la Norma ISO 19115-1):

- a) Una serie de conjuntos de datos;
- b) Un conjunto de datos;
- c) Un subconjunto de datos definido por una o más de las siguientes características:
 - 1) Tipos de ítems (conjuntos de tipos, atributos, operaciones o relaciones entre objetos geográficos),
 - 2) Ítems específicos (conjuntos de tipos, atributos, operaciones o relaciones entre objetos geográficos),
 - 3) Una extensión geográfica,
 - 4) Una extensión temporal (el período de tiempo de referencia y su exactitud);
- d) Una cobertura

8.3 Elementos de la calidad de los datos

8.3.1 Generalidades

Un elemento de la calidad de los datos es una componente que describe un cierto aspecto de la calidad de los datos geográficos. En este documento, estos elementos se organizan en grupos de elementos de calidad que se definen en los apartados del 8.3.2 a 8.3.7. Las relaciones entre varios aspectos de la calidad y sus correspondientes elementos de la calidad se presentan en la figura 4, donde los grupos de elementos se representan mediante clases abstractas y los elementos mediante sus subclases.

NOTA: Las aplicaciones del modelo conceptual de la calidad de los datos presentado en la figura 4 varían dependiendo de la herramienta utilizada o de las preferencias organizativas. El asesoramiento sobre las mejores prácticas de aplicación está fuera del objeto de este documento.

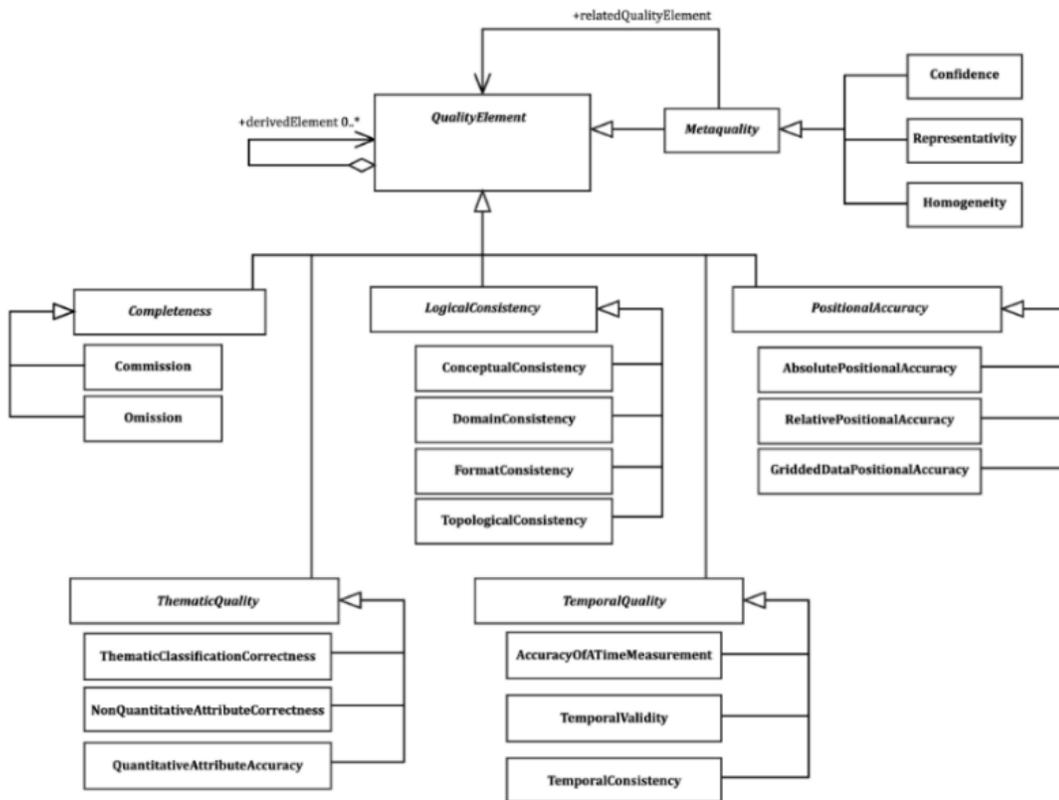


Figura 4 – Visión de los elementos de la calidad de los datos

Permiso 1:	https://standards.iso/211.org/19157-1/1/per/permittedContent/additionalQualityElement
	Si los elementos de la calidad definidos en este documento no son suficientes para expresar la calidad del conjunto de datos, se

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

	pueden definir y utilizar nuevos elementos según los requisitos especificados en el apartado 8.4.
--	---------------------------------------------------------------------------------------------------

8.3.2 Completitud

La completitud se define como el grado en que un conjunto de datos (o una serie de conjuntos de datos) que representa un objeto geográfico (o varios) tiene, en el momento que se creó, valores para todos los atributos esperados e instancias de entidad relacionadas (ISO 8000-2; ISO/IEC 25012). Por lo tanto, la completitud se refiere a la presencia o ausencia de objetos geográficos de objetos geográficos, sus atributos y relaciones. En este documento, dos elementos (y sus medidas asociados) definen la completitud:

- Coherencia conceptual: adherencia a las reglas del modelo conceptual;
- Omisión, datos ausentes de un conjunto de datos.

8.3.3 Coherencia lógica

La coherencia lógica se define como el grado de adherencia a las reglas lógicas de la estructura de los datos, de los atributos y de las relaciones (la estructura de los datos puede ser conceptual, lógica o física). Si estas reglas lógicas se documentan en otro lugar (por ejemplo, en unas especificaciones de producto de datos), entonces se debería referenciar la fuente (por ejemplo, en la evaluación de la calidad de los datos). Este documento especifica la coherencia lógica con los elementos de la calidad de los datos (y sus medidas asociadas):

- Coherencia conceptual: adherencia a las reglas del modelo conceptual;
- Coherencia de dominio: adherencia de los valores a su dominio;
- Coherencia de formato: grado en el que los datos se almacenan de acuerdo con la estructura física del conjunto de datos;
- Coherencia topológica: corrección de las características topológicas codificadas explícitamente.

8.3.4 Exactitud posicional

La exactitud posicional se relaciona con la exactitud de medida (Guía ISO/IEC 98-3; Guía ISO/IEC 99) y representa el grado de concordancia entre una posición medida de los objetos geográficos y una posición aceptada como verdadera en un determinado sistema de referencia espacial. Dependiendo del ámbito y del tipo de sistema de referencia, este documento especifica los siguientes tres elementos de la calidad de los datos para expresar la exactitud posicional:

- Exactitud absoluta o externa: proximidad de los valores consignados de las

coordenadas a los valores aceptados como verdaderos en un sistema de referencia de coordenadas estandarizado;

- Exactitud relativa o interna: proximidad de las posiciones relativas de los objetos geográficos de un conjunto de datos relacionado a sus respectivas posiciones relativas aceptadas como verdaderas en un sistema de referencia de coordenadas local;
- Exactitud posicional de datos en malla: proximidad de los valores de posición de los datos en estructura de malla regular a los valores aceptados como verdaderos.

La exactitud posicional es una estimación de la incertidumbre de los resultados de medición (Norma ISO 19116) y por tanto debería expresarse con medidas de incertidumbre.

8.3.5 Exactitud posicional

La exactitud posicional se relaciona con la exactitud de medida (Guía ISO/IEC 98-3; Guía ISO/IEC 99) y representa el grado de concordancia entre una posición medida de los objetos geográficos y una posición aceptada como verdadera en un determinado sistema de referencia espacial. Dependiendo del ámbito y del tipo de sistema de referencia, este documento especifica los siguientes tres elementos de la calidad de los datos para expresar la exactitud posicional:

- Exactitud absoluta o externa: proximidad de los valores consignados de las coordenadas a los valores aceptados como verdaderos en un sistema de referencia de coordenadas estandarizado;
- Exactitud relativa o interna: proximidad de las posiciones relativas de los objetos geográficos de un conjunto de datos relacionado a sus respectivas posiciones relativas
- Exactitud posicional de datos en malla: proximidad de los valores de posición de los datos en estructura de malla regular a los valores aceptados como verdaderos.

La exactitud posicional es una estimación de la incertidumbre de los resultados de medición (Norma ISO 19116) y por tanto debería expresarse con medidas de incertidumbre.

8.3.6 Calidad temática

La calidad temporal se define como la calidad de los atributos y de las relaciones temporales de los objetos geográficos. Puede describirse por los siguientes elementos de la calidad de los datos (y sus medidas asociadas):

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Corrección de la clasificación: comparación de las clases asignadas a los objetivos geográficos o a sus atributos, frente un universo de discurso (por ejemplo, la verdad terreno o unos datos de referencia);
- Corrección de atributos no cuantitativos: medida de si un atributo no cuantitativo es correcto o incorrecto;
- Exactitud de atributos cuantitativos: proximidad del valor de un atributo cuantitativo al valor verdadero o al aceptado como tal.

8.3.7. Elementos de metacalidad

Los elementos de metacalidad son un conjunto de declaraciones cuantitativas y cualitativas sobre una evaluación de la calidad y su resultado. El conocimiento de la calidad y la adecuación del método de evaluación, la medida aplicada y el resultado ofrecido pueden ser de la misma importancia que el resultado en sí mismo.

Véase el apartado D.5.3 para un ejemplo de evaluación de la metacalidad.

La metacalidad debería describirse utilizando los siguientes elementos.

- Confianza: fiabilidad de un resultado de la calidad de los datos.

NOTA 1: Pueden obtenerse cifras cuantitativas de confianza mediante parámetros estadísticos tales como la desviación típica o un intervalo de confianza a un nivel de confianza dado.

EJEMPLO: La confianza procede principalmente del método usado y su fiabilidad, además de, en menor medida, a la población de que se trate.

- Representatividad: grado en que la muestra utilizada ha producido un resultado que es representativo de los datos pertenecientes al ámbito de calidad de los datos.

NOTA 2: Un método estadístico basado en el muestreo podría considerarse como fiable en tanto que método global cuando se cubren todas las zonas geográficas y períodos de tiempo afectados y la población es suficientemente grande. No solo es crucial el tamaño de la muestra sino también la medida en que representa satisfactoriamente el estado real de los datos. Véase también el apartado 10.2.2 y el anexo E.

- Homogeneidad: uniformidad esperada o sometida a ensayo de los resultados obtenidos para una evaluación de la calidad de los datos.

NOTA 3: La homogeneidad consiste en la comparación de los resultados de la evaluación de varios segmentos de un conjunto de datos global. Por ejemplo, esta comparación puede expresarse usando valores eficaces de error (como en el apartado E.5.3). En el caso de un proceso general, la homogeneidad no puede evaluarse de tratarse de un resultado global.

NOTA 4: Estas pruebas se realizan a menudo cuando los datos han capturado diferentes operadores, dependiendo de la zona o de la fecha de adquisición.

8.4 Ampliación del modelo de información de la calidad de los datos

Este documento reconoce que para muchos fines de campos específicos es necesario o conveniente extender el modelo normalizado de información de la calidad de los datos tal y como está definido en el apartado 8.3. La ampliación incluye la incorporación de elementos de calidad de los datos (o una de sus especializaciones tal y como se definen en el apartado 8.3) y medidas de calidad de los datos.

EJEMPLO: Los productores de datos pueden querer proporcionar medios para que sus usuarios registren comentarios sobre un conjunto de datos. Por ejemplo, se puede definir un modelo de información para esos comentarios (es decir, nuevos elementos de calidad) con elementos de la referencia [36].

Requisito 2:	https://standards.iso/211.org/19157-1/1/req/content/additionalQualityElement
	<p>Un nuevo elemento de la calidad debe definirse como una especialización de un QualityElement, (véase 8.3) de acuerdo con las reglas para las ampliaciones de metadatos (Norma ISO 19115-1:2014, anexo C) y las reglas para perfiles de esquemas normalizados para un dominio (Norma ISO 19109:2015, apartado 8.3).</p> <p>NOTA: Si procede, los elementos definidos en otras normas internacionales (por ejemplo, la Norma ISO/IEC 25000) se pueden utilizar para ampliar el conjunto de componentes de la calidad definido en el capítulo 8.</p>

8.5 Descriptores de los elementos de la calidad de datos

8.5.1 Ampliación del modelo de información de la calidad de los datos

Un elemento de la calidad de datos se describe con una referencia a una medida de la calidad, un método de evaluación, un resultado de la calidad y un elemento de la metacalidad. Estos descriptores se muestran en la figura 5 y se describen en los apartados 8.5.2 a 8.5.5.

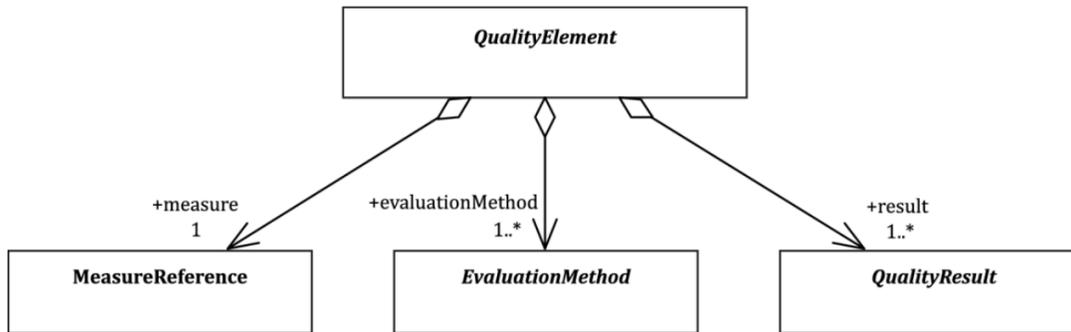


Figura 5 – Descriptores de un elemento de la calidad de los datos

8.5.2 Referencia de medida

Los elementos de la calidad de los datos hacen referencia a las medidas por medio de una referencia de medida (véase la figura 6), proporcionando el identificador de una medida detallada al completo en otro lugar.

NOTA: La descripción completa puede encontrarse en un registro o catálogo de medidas, que puede formar parte de unas especificaciones de producto de datos o de un informe de evaluación de la calidad.

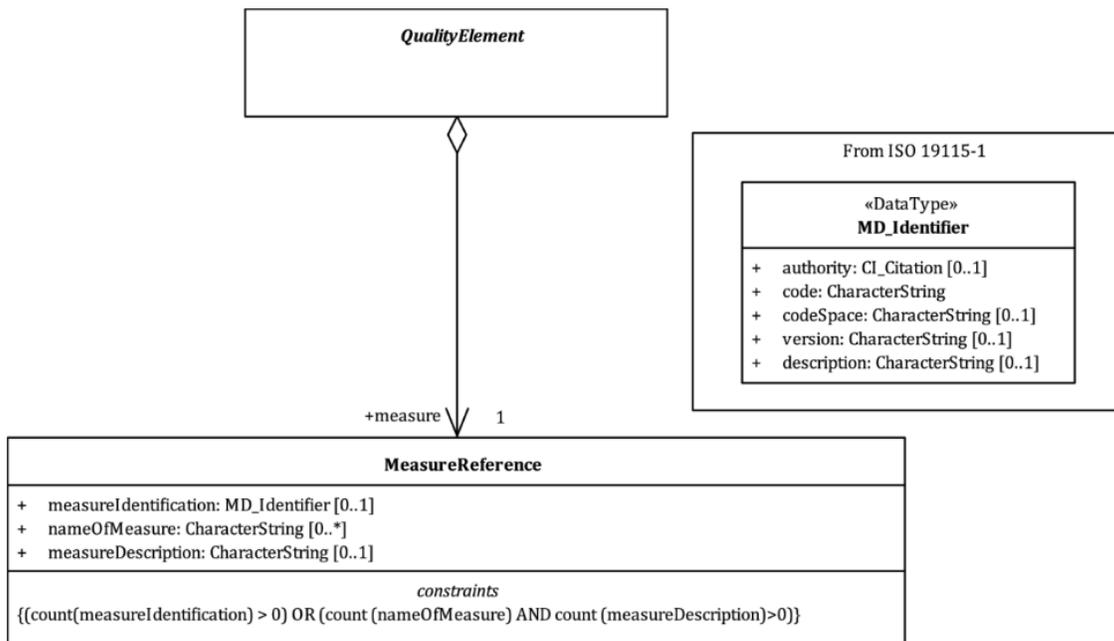


Figura 6 – Referencia de medida de la calidad de los datos

Las medidas de calidad de los datos se describen más a fondo en el capítulo 0 de este documento,

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

EJEMPLO: El porcentaje de valores correctos de un atributo.

Este documento reconoce que la calidad de un conjunto de datos se mide empleando diversos métodos. Una única medida de la calidad puede ser insuficiente para evaluar, de forma completa, la calidad de los datos especificados por un ámbito y también lo podría ser frente a todas las posibles utilidades del conjunto de datos. Una combinación de medidas de calidad puede proporcionar información útil y por ello se pueden ofrecer múltiples medidas para los datos especificados por un ámbito. En este caso el informe sobre la calidad de los datos debería incluir una instancia de *QualityElement* para cada medida aplicada.

8.5.3 Método de evaluación

Un método de evaluación de la calidad de los datos describe aquellos procedimientos y métodos que se aplican a los datos geográficos para obtener un resultado de la calidad de los datos; véase la figura 7. A menudo se usan diferentes evaluaciones para los diferentes elementos de la calidad de los datos.

Un método de evaluación de la calidad de los datos se usa para describir, o referenciar documentación que describa, la metodología empleada para aplicar una medida de la calidad a los datos especificados por su ámbito de calidad.

NOTA: La evaluación de la calidad de los datos se describe más a fondo en el capítulo 10.

EJEMPLO: Ejemplos de documentación son la especificaciones de producto de datos (por ejemplo, la Norma ISO 19131), los artículos publicados o las normas aceptadas por la industria.

Recomendación 1:	https://standards.iso/211.org/19157/-1/1/rec/optionalContent/dateTime
	Debería incluirse una fecha o un rango de fechas (de acuerdo con la Norma ISO 19108) para cada evaluación. Si la evaluación se realizó en fechas no consecutivas, debería incluirse cada fecha individual.

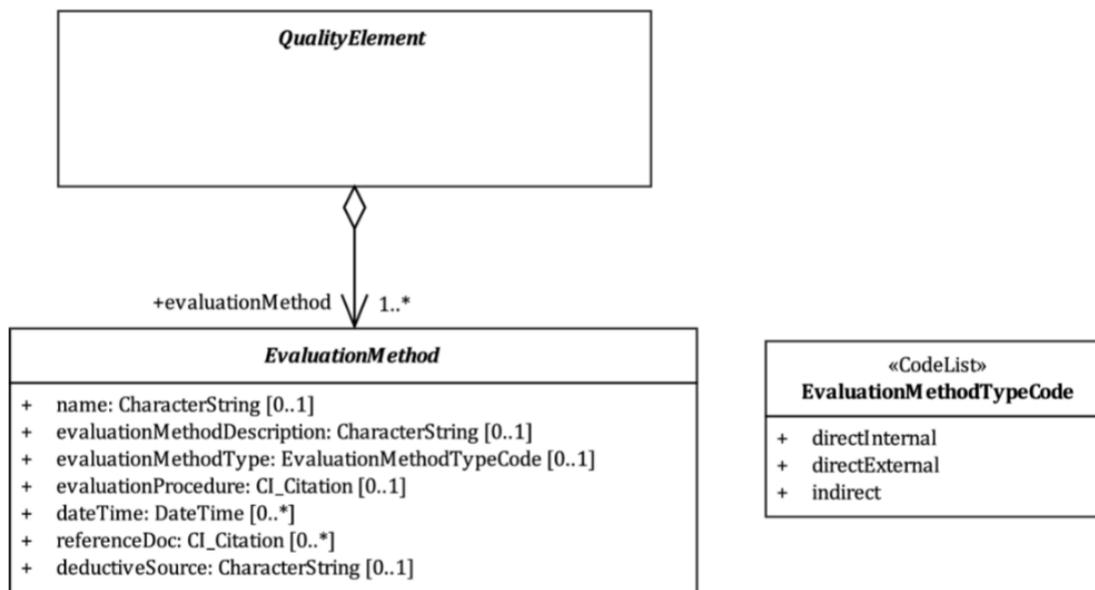


Figura 7 – Método de evaluación de la calidad de los datos

8.5.4 Resultado de la calidad

8.5.4.1 Generalidades

Se proporciona un resultado de la calidad de los datos para cada elemento de la calidad de los datos. Puede ser un resultado cuantitativo, de conformidad, descriptivo o en cobertura; véase también la figura 8.

NOTA: Pueden proporcionarse diferentes tipos de resultados para el mismo elemento de la calidad de los datos.

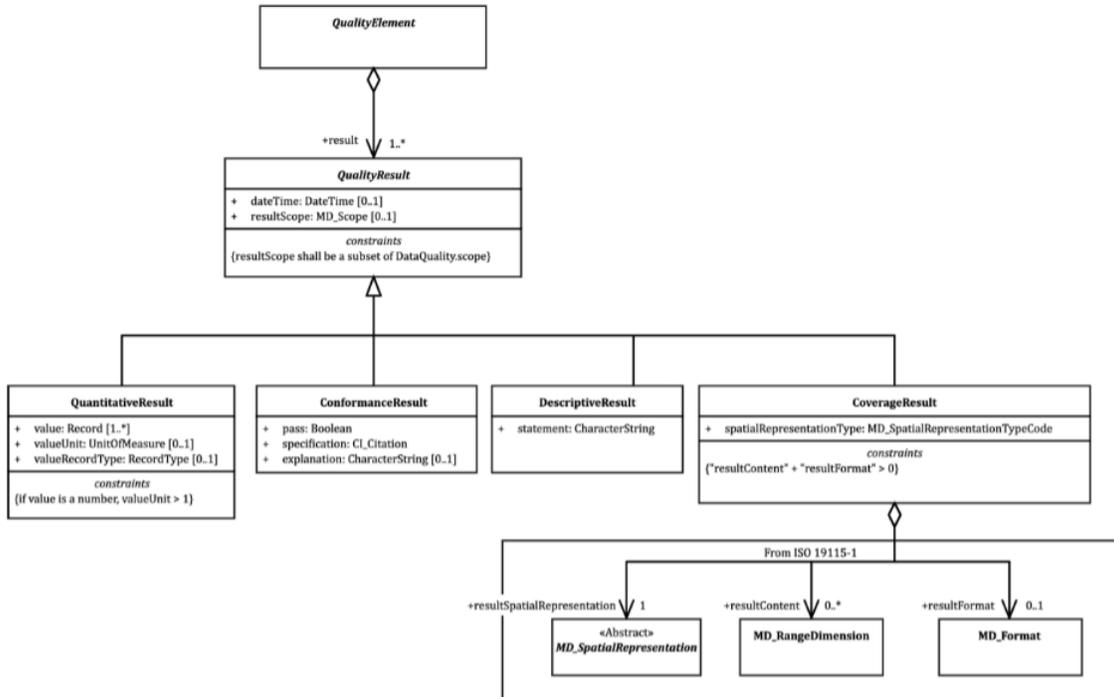


Figura 8 – Resultado de la calidad de los datos

La calidad difiere con la frecuencia entre diferentes partes del conjunto de datos que se evalúa. Por lo tanto, pueden aplicarse varias evaluaciones para el mismo elemento de la calidad de los datos para describir su calidad cuantitativa de forma más completa y con mayor detalle. Pueden usarse varios resultados con ámbitos individuales para evitar la repetición de las descripciones de la medida y del procedimiento de evaluación en varias instancias de elemento de la calidad de los datos (*QualityElement*).

EJEMPLO: Un conjunto de datos contiene objetos geográficos de tipo idéntico pero cuyas posiciones se han establecido con varios métodos dando diferentes exactitudes posicionales. No obstante, se ha aplicado el mismo método de evaluación de la calidad y la misma medida para todo el conjunto de datos y se proporcionan diferentes resultados dependiendo del método de adquisición. En este caso, puede ser deseable disponer de varios resultados con ámbitos individuales (el área cubierta por cada método de adquisición) y un único ámbito de la calidad de los datos (el conjunto de datos).

8.5.4.2 Resultado cuantitativo

Un resultado cuantitativo puede ser un valor individual o valores múltiples, dependiendo de los valores de los atributos *valueType* y *valueStructure* definidos en la descripción de la medida aplicada.

El atributo *valueRecordType* se usa para describir cómo los atributos *valueType* y *valueStructure* definidos en la medida se implementan para proporcionar el valor del resultado cuantitativo.

NOTA: El atributo *valueRecordType* es el tipo *RecordType*, que es un tipo de datos genérico definido en la Norma ISO 19103. Su valor cambia dependiendo de qué solución de implementación se ha usado para proporcionar el resultado cualitativo. Un ejemplo de aplicación de lenguaje extensible de marcado (XML) para *recordType* se proporciona en la Especificación Técnica ISO/TS 19139-1.

EJEMPLO 1: Usando una implementación XML: ejemplo simple: *value* = 5, *valueRecordType* = *gco:Integer*, *valueUnit* = "metre".

EJEMPLO 2: Dentro de la descripción de la medida, el *valueType* es un entero, el *valueStructure* es una matriz $n \times n$ (cuadrada). El atributo *value* del resultado cuantitativo proporciona el propio resultado de la matriz, dentro de una codificación numérica que usa un tipo XML particular denominado *MatrixType*. El atributo *valueRecordType* proporciona la descripción del tipo *MatrixType* en XML. Si se usa otra codificación, el atributo *valueRecordType* cambiará para proporcionar la descripción del tipo *Matrix* en la otra codificación y en consecuencia cambiará la implementación del atributo *value*, pero no lo hará el valor en sí mismo.

EJEMPLO 3: La medida "Índice de ítems en exceso" se usa para evaluar el número de ítems en exceso en el conjunto de datos en relación con el número de ítems que se espera que estén presentes. El valor del resultado cuantitativo es del tipo valor Real. La unidad del valor se usa en este caso para mostrar que el valor es un porcentaje, el valor se ha multiplicado por 100. En este ejemplo la unidad del valor es "%".

8.5.4.3. Resultado de conformidad

Un resultado de conformidad se deriva de la comparación del valor, o conjunto de valores, obtenido de la aplicación de una medida de calidad a unos datos especificados por un ámbito frente a un nivel de conformidad de la calidad especificado (por ejemplo, según lo registrado en una especificaciones de producto de datos conformes con la Norma ISO 19131).

Cuando se define un nivel de conformidad de la calidad, el resultado obtenido se compara con este para evaluar si la calidad de los datos se adecua al nivel de calidad especificado.

Se puede proporcionar un resultado de conformidad para cada medida. El nivel de

conformidad de la calidad usualmente está especificado en al documentación de referencia adecuada, como son las especificaciones de producto de datos (por ejemplo, unas especificaciones de producto de datos conformes con la Norma ISO 19131) o una especificación de requisitos definidos por el usuario. Si se evalúa la conformidad, debe indicarse la documentación de referencia que sea relevante y debe especificarse el nivel de conformidad de la calidad utilizado.

El resultado de conformidad no tiene que usarse para especificar el nivel de conformidad de la calidad de unas especificaciones de producto de datos conformes con la Norma ISO 19131.

Recomendación 2:	https://standards.iso211.org/19157-1/1/rec/optionalContent/identifier
	<p>Se debería proporcionar un <i>identifier</i> para el atributo <i>specification</i> de <i>ConformanceResult</i></p> <p>NOTA: <i>Specification</i> es el tipo de <i>CI_Citation</i>, en el que, según lo definido por la Norma ISO 19115-1, <i>identifier</i> es opcional</p>

Permiso 2:	https://standards.iso211.org/19157-1/1/per/permittedContent/conformanceResult
	<p>Si la evaluación se ha realizado frente a niveles de conformidad procedentes de diferentes fuentes, puede proporcionarse más de un resultado de conformidad <i>ConformanceResult</i> para la misma medida</p>

8.5.4.4. Resultado descriptivo

En algunos casos (por ejemplo, con observaciones temáticas y geo-científicas), no es posible obtener un resultado cuantitativo para un elemento de la calidad de los datos. En ese caso puede expresarse una evaluación subjetiva del elemento con una declaración textual como un resultado descriptivo de la calidad.

EJEMPLO: La exactitud posicional relativa entre un objeto geográfico

geológico y un objeto cercano de un mapa base (carreteras, ríos, lagos, etc.) es mayor que la exactitud posicional absoluta del objeto geográfico geológico.

Este resultado descriptivo también puede utilizarse para proporcionar una breve descripción resumida del resultado de la evaluación de la calidad, para acompañar al resultado cuantitativo completo o reemplazarlo si no se puede proporcionar ningún resultado cuantitativo.

8.4.4.5 Resultado en cobertura

Un resultado de la calidad puede proporcionarse como información en una cobertura. En la figura 8 se proporciona un modelo UML para un resultado en cobertura.

8.5.5. Descriptores de un elemento de metacalidad

Un elemento de metacalidad se describe mediante los mismos descriptores que un elemento de calidad: medida, método de evaluación y resultado (véase la figura 9).

NOTA: El elemento de calidad relacionado (*relatedQualityElement*) es el elemento sobre el que se aplica el elemento de metacalidad.

Véase el apartado E.5.3 para un ejemplo de evaluación de metacalidad.

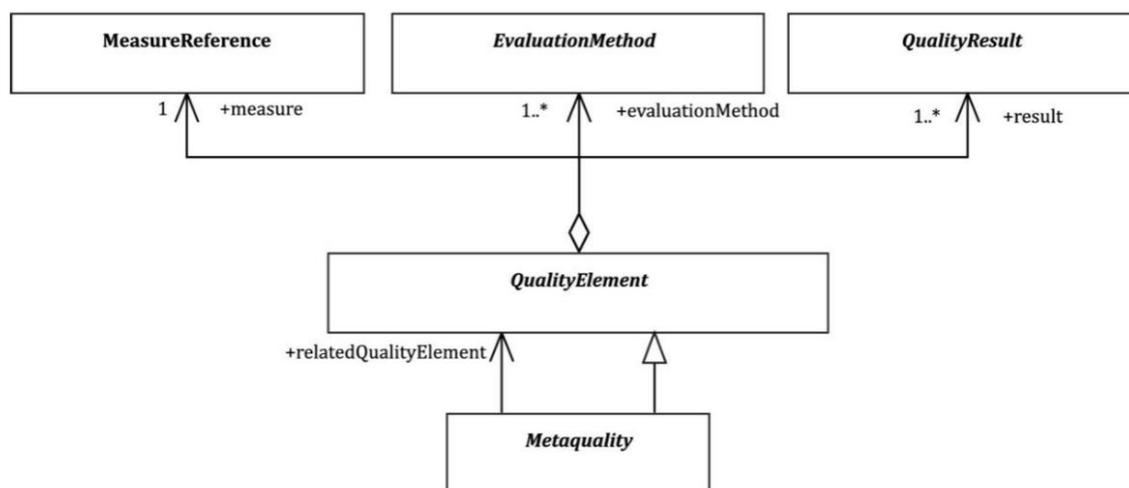


Figura 9 – Descriptores de metacalidad

9. MEDIDAS DE CALIDAD DE LOS DATOS

9.1. Generalidades

Para facilitar las comparaciones entre conjuntos de datos, las evaluaciones y los informes de la calidad de los datos (metadatos o informes de evaluación de la calidad) han de expresarse de forma comparable y es necesario que las medidas

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

de calidad que se han usado se interpreten de la misma manera.

Requisito 3:	https://standards.iso/211.org/19157-1/1/req/content/qualityMeasure
	<p>Quando sea posible, deben usarse medidas definidas en un registro de medidas de calidad para información geográfica en la evaluaciones de calidad de los datos y para representar el resultado de una evaluación de la calidad en informes de evaluación de la calidad y en metadatos.</p> <p>NOTA 1: Si no es aplicable ninguna de las medidas definidas en un registro de medidas de calidad para información geográfica, pueden utilizarse medidas adicionales definidas por el usuario.</p> <p>NOTA 2: Si procede, los elementos definidos en otras normas internacionales (por ejemplo, la Norma ISO/IEC 25000) se pueden utilizar para ampliar el modelo de la calidad.</p>

9.2 Medidas normalizadas de la calidad de los datos

9.2.1 Generalidades

La estructura de las medidas normalizadas de la calidad se presenta en la figura 10.

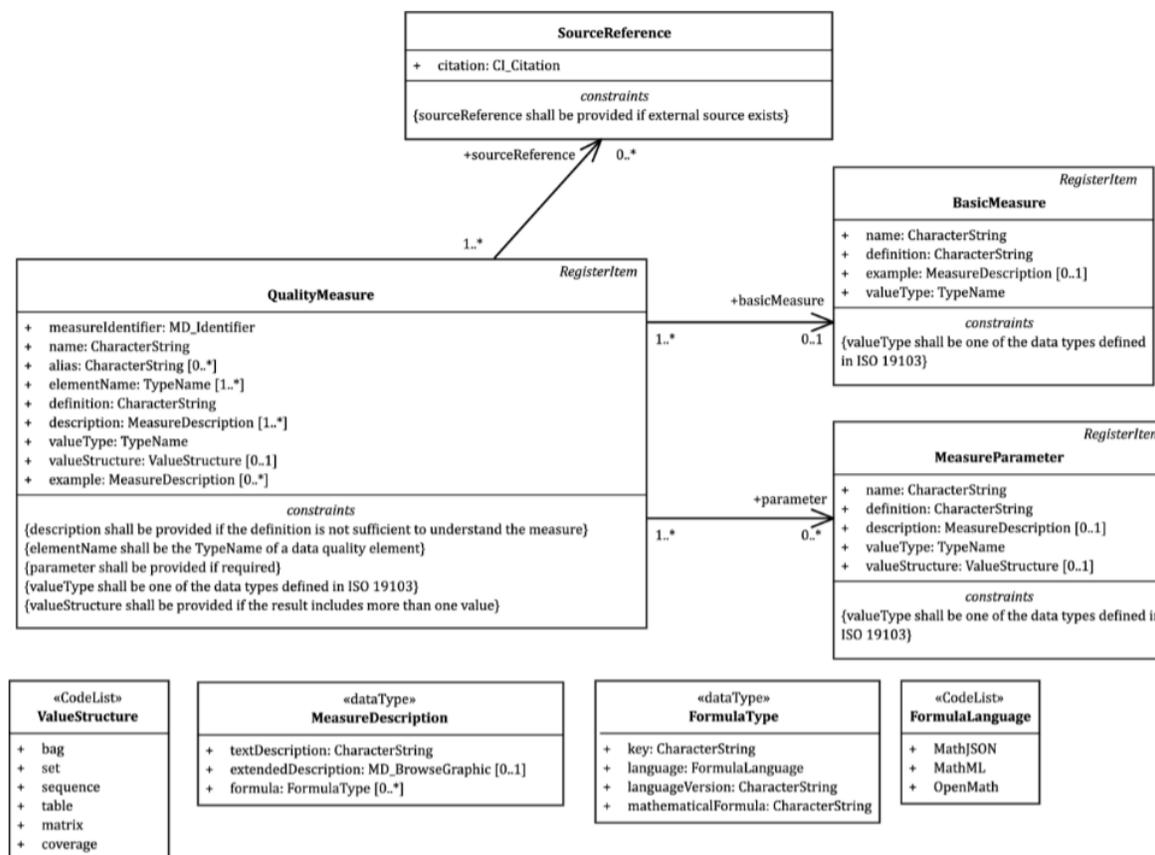


Figura 10 – Medidas de calidad de los datos

Cada medida de calidad de los datos se describe mediante los siguientes componentes:

- Identificador de la medida (9.2.2);
- Nombre (9.2.3);
- Sobrenombre (9.2.4);
- Medida básica (9.2.6);
- Definición (9.2.7);
- Descripción (9.2.8);
- Parámetro (9.2.9);
- Tipo de valor (9.2.10);
- Estructura del calor (9.2.11);

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Fuente de referencia (9.2.12);
- Ejemplo (9.2.13).

Para cada elemento de calidad pueden definirse múltiples medidas, cuya elección dependerá del tipo de datos y su uso previsto. Se proporciona una lista de medidas normalizadas en un registro de medidas de calidad de los datos conforme con la Norma ISO 19135-1.

9.2.2 Identificador de la medida

El identificador es un valor que identifica de manera única una medida dentro de un espacio de nombres.

NOTA: Este identificador posibilita las referencias a la medida de la calidad en los elementos de la calidad (véase 8.5.2)

9.2.3 Nombre

El nombre es el nombre de la medida.

NOTA: Si la medida ya dispone de un nombre de uso común, se utiliza este nombre. Si no existe ningún nombre, se elige uno que refleje la naturaleza de la medida.

9.2.4 Sobrenombre

El sobrenombre es otro nombre reconocido para la misma medida de la calidad. Puede ser otro nombre de uso común, una abreviatura o un nombre corto. Se puede proporcionar más de un sobrenombre.

9.2.5 Nombre del elemento

El nombre del elemento es el nombre del elemento de calidad de los datos (véanse 8.3 y 8.4) al que se aplica la medida. Se puede proporcionar más de un nombre.

9.2.6 Medida básica

Las medidas básicas se definen por *name*, *definition*, *example* y *valueType* (véase la figura 10).

Si procede, las medidas básicas también deberían usarse para crear nuevas medidas. Por ejemplo, medidas para informar sobre las superficies planas sin cerrar u otras medidas que dependan de la aplicación.

Recomendación 3:	https://standards.iso/211.org/19157/-1/1/rec/optionalContent/additionalQualityMeasure
---------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>Siempre que sea posible, debería usarse <i>BasicMeasure</i> como base para la definición de una nueva medida de la calidad.</p> <p>NOTA: Diversas medidas se basan en el recuento de ítems erróneos. También existen varias medidas relacionadas con la incertidumbre de valores numéricos. Para evitar la repetición, se definen, en cuanto a medidas básicas, los métodos más comunes para construir medidas de recuento, así como medidas estadísticas generales para variables aleatorias de una y dos dimensiones.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

9.2.7 Definición

La definición es una representación de un concepto mediante un enunciado descriptivo que permite diferenciarlo de otros conceptos relacionados.

NOTA: Si la medida se deriva de una medida básica, la definición se basa en la de la medida básica y se especializa para esa medida.

9.2.8 Descripción

La descripción de la medida es la descripción que incluye los métodos de cálculo, con toda la formulación y las ilustraciones necesarias para determinar el resultado de la aplicación de la medida.

Si la medida utiliza el concepto de error, debería indicarse cuándo un ítem se clasifica como incorrecto. Esto ocurre cuando la calidad solo puede consignarse como correcta o incorrecta.

<p>Recomendación 4:</p>	<p>https://standards.iso/211.org/19157-1/1/rec/optionalContent/description</p>
	<p>Si la medida utiliza el concepto de error para informar sobre ítems correctos o incorrectos, <i>description</i> debería contener una explicación sobre cuándo se clasifica un ítem como incorrecto.</p>

9.2.9 Parámetro

El parámetro de una medida es una variable auxiliar utilizada por la medida.

9.2.10 Tipo de valor

El tipo de valor es el tipo de datos usado para informar del resultado de la medida. Se deben usar los tipos de datos definidos en la Norma ISO 19103.

9.2.11 Estructura del valor

Un resultado puede estar formado por valores múltiples. En tales casos, el resultado debe estructurarse usando la estructura del valor como se indica en el apartado C.3.3.

9.2.12 Fuente de referencia

La fuente de referencia es la cita de la documentación de la medida.

9.2.13 Ejemplo

El ejemplo es un ejemplo de aplicación de la medida o del resultado obtenido con ella. Se puede proporcionar más de un ejemplo.

9.3 Medidas de calidad de los datos definidas por el usuario

Debido a la naturaleza de los datos geográficos y de su calidad, la lista de medidas de calidad normalizadas no puede completarse. Pueden existir casos donde el usuario de este documento tenga que concebir otras medidas de calidad de los datos. Cuando sea posible, esas medidas deben definirse usando las medidas básicas de calidad de los datos y la estructura indicada en el capítulo 9. En el anexo C de la Norma ISO 19160-3:2020 se pueden encontrar ejemplos adicionales de medidas de calidad de los datos.

Requisito 4:	https://standards.iso/211.org/19157-1/1/req/content/additionalQualityMeasure
	Una nueva medida de la calidad debe definirse como una especialización de <i>QualityMeasure</i> , o de uno de sus subtipos (véase el capítulo 9) según las reglas para ampliaciones de metadatos (Norma ISO 19115-1:2014, anexo C) y las reglas para perfiles de esquemas normalizados para un dominio (Norma ISO 19109:2015, apartado 8.3).

	<p>Si la medida se deriva de una medida básica, esta debe referenciarse.</p> <p>Si se adopta un ítem de una fuente externa, debe proporcionarse una referencia <i>MeasureParameter</i>.</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

10. EVALUACIÓN DE LA CALIDAD DE LOS DATOS

10.1 El proceso de evaluación de la calidad de los datos

10.1.1. Introducción

Los procesos de evaluación de la calidad se usan en diferentes fases del ciclo de vida de un producto, teniendo diferentes objetivos en cada una de ellas. Las fases del ciclo de vida aquí consideradas son: especificación, producción, entrega, uso y actualización.

El proceso de evaluación de la calidad de los datos es una secuencia de pasos orientados a producir un resultado de la calidad de los datos.

10.1.2. Flujo del proceso

El proceso de evaluación de la calidad es una secuencia de pasos que se siguen para producir un resultado de la evaluación de la calidad. La figura 11 ilustra un posible flujo de trabajo para evaluar la calidad de los datos; véase también el anexo D para una descripción de los conceptos para evaluar e informar sobre la calidad de los datos.

<p>Recomendación 5:</p>	<p>https://standards.iso.org/standards.iso.org/19157-1/1/rec/content/qualityEvaluationProcess</p>
	<p>Cuando el conjunto de datos geográficos evaluado es heterogéneo y la calidad se ha especificado para distintas partes, deberían realizarse evaluaciones independientes para las distintas partes del conjunto de datos.</p>

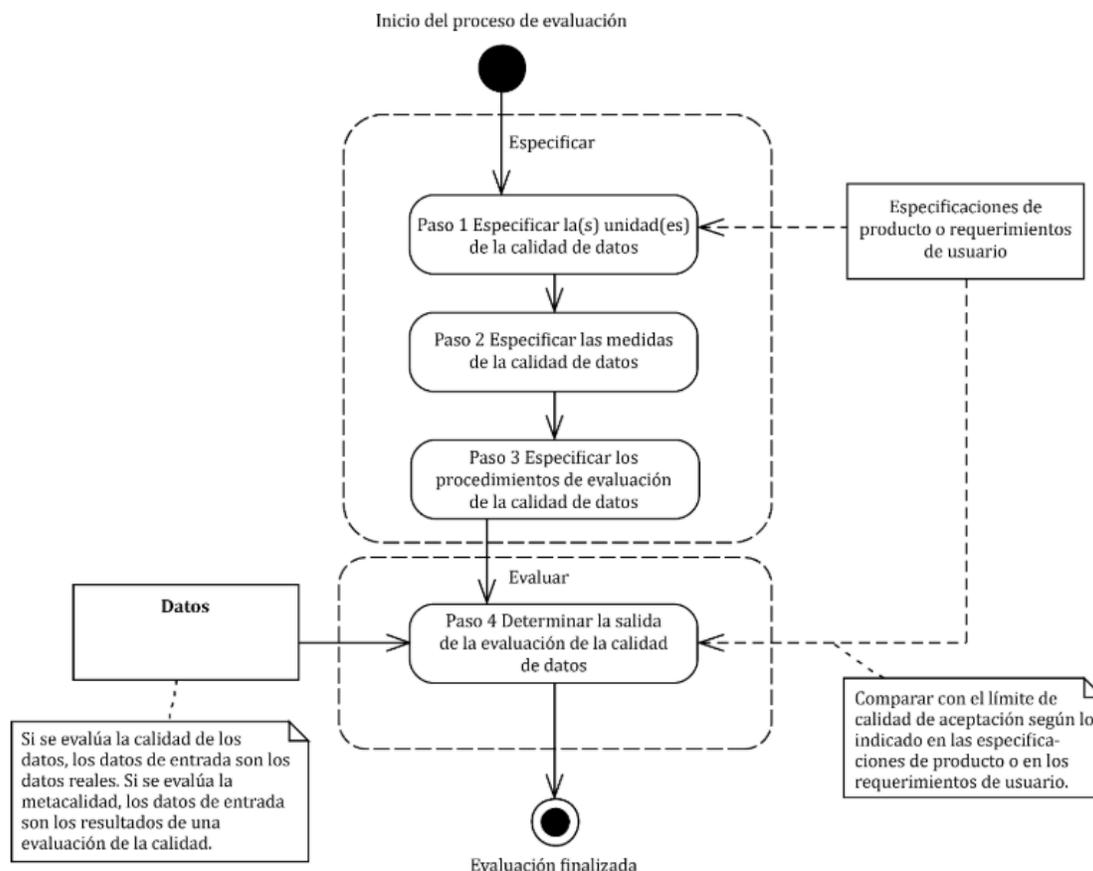


Figura 11 – Evaluación de la calidad de los datos

10.1.3 Pasos del proceso

En la tabla 7 se especifican los pasos del proceso que intervienen en la evaluación de la calidad, como se ilustra en la figura 11.

Tabla 7 – Pasos del proceso

Paso del proceso	Acción	Descripción
1	Especificar la(s) unidad(es) de la calidad de los datos	Una unidad de la calidad de los datos se compone de un ámbito y de elemento(s) de la calidad; véase el apartado 8.2. Deberían utilizarse todos los elementos pertinentes para los datos cuya calidad se tiene que describir. NOTA En el apartado 8.3 se describen los elementos de la calidad a evaluar y en el anexo D se proporcionan pautas para el uso de estos elementos. Según se describe en el apartado 8.4, se pueden definir elementos adicionales y utilizarlos en la evaluación de la calidad.
2	Especificar las medidas de calidad de los datos	Debería especificarse una medida para cada elemento de la calidad. Si no se identifica ninguna medida, puede proporcionarse un resultado descriptivo.
3	Especificar los procedimientos de evaluación de la calidad de los datos	Un procedimiento de evaluación de la calidad de los datos consiste en la aplicación de uno o más métodos de evaluación. Incluidos: – método de evaluación estandarizado que incluye niveles de conformidad; – estrategia de muestreo.
4	Determinar el producto de la evaluación de la calidad de los datos	El producto de aplicar la evaluación es un resultado. La conformidad con las especificaciones del producto de datos o con los requerimientos del usuario es parte del producto de la evaluación de la calidad. El producto se presenta como metadatos (por ejemplo, en MD_Usage definido en la Norma ISO 19115-1) y como informe de evaluación de la calidad.

La evaluación de la metacalidad puede realizarse tras la obtención del resultado de la evaluación de la calidad. El flujo de trabajo descrito anteriormente también es aplicable para evaluar la metacalidad, con los siguientes pasos del proceso: especificar el elemento de la calidad y la evaluación de la calidad para los que se va a evaluar la metacalidad, especificar una medida y un método de evaluación y determinar el resultado de la evaluación de la metacalidad.

10.2. Métodos de evaluación de la calidad de los datos

10.2.1 Clasificación de los métodos de evaluación de la calidad de los datos

Un procedimiento de evaluación de la calidad de los datos comprende uno o más métodos de evaluación de la calidad de los datos. Los métodos de evaluación de la calidad de los datos pueden dividirse en dos clases principales: directos e indirectos. Los métodos directos determinen la calidad mediante la comparación de los datos con información de referencia interna o externa. Los métodos indirectos infieren o estiman la calidad usando información sobre los propios datos, como el linaje.

Recomendación	https://standards.iso/211.org/19157/-1/1/rec/optionalContent/evaluationMethodT
---------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

n 6:	ype
	Se deberían usar preferentemente métodos de evaluación directos (<i>directInternal</i> o <i>directExternal</i>) en lugar de evaluaciones indirectas

Los métodos de evaluación directos se pueden a su vez subclasificar en internos o externos, atendiendo a la fuente de información necesaria para realizar la evaluación. En la figura 12 se muestran las clases usadas para describir los métodos de evaluación.

NOTA: El linaje se describe en la norma 19115-1.

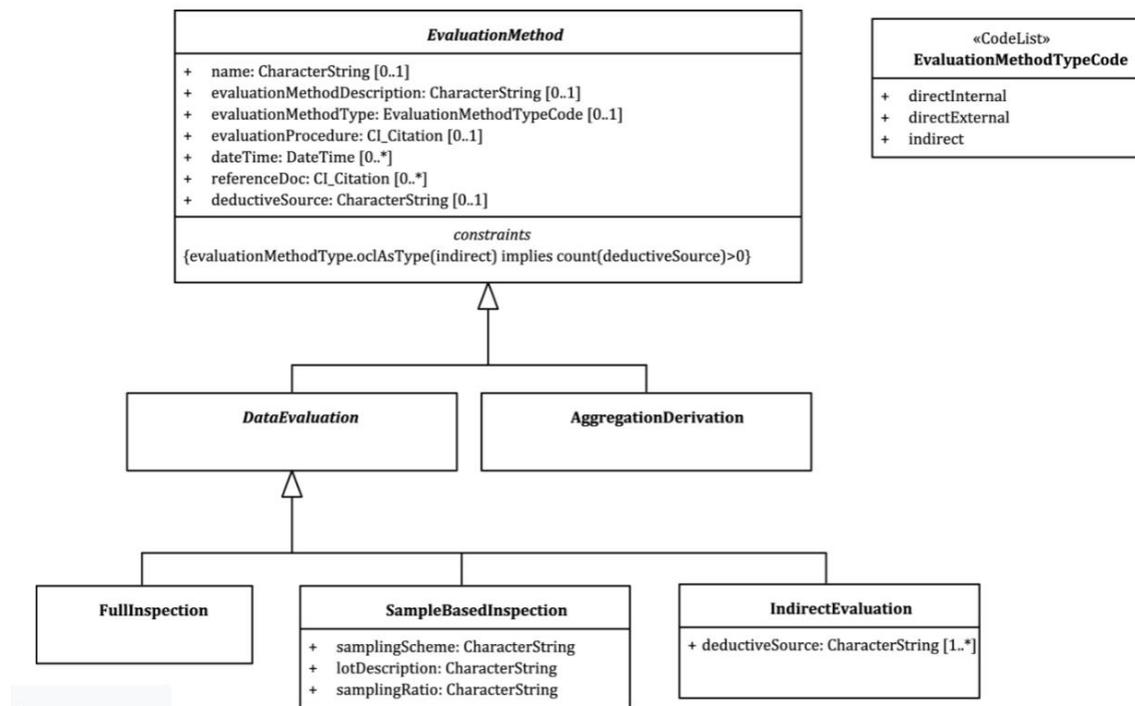


Figura 12 – Métodos de evaluación de la calidad de datos

10.2.2 Evaluación directa

Un método de evaluación directo es un método de evaluación de la calidad de un conjunto de datos basado en la inspección de los ítems del conjunto de datos.

Los métodos de evaluación directos pueden clasificarse como internos o externos.

La evaluación de la calidad directa interna utiliza solo datos que se encuentran en el propio conjunto de datos que se está evaluando. La evaluación de la calidad

directa externa requiere datos de referencia externos al conjunto de datos que se está evaluando.

NOTA: Los datos de referencia son datos aceptados como la representación del universo de discurso.

Para los métodos de evaluación directos, tanto externos como internos, puede usarse uno de los siguientes métodos de inspección:

- Inspección completa: la evaluación de la calidad comprueba cada ítem especificado por el ámbito de la calidad de los datos.
- Inspección por muestreo: la evaluación de la calidad se realiza en subconjuntos de los datos geográficos definidos por el ámbito de la calidad de los datos.

Si la evaluación se basa en la inspección por muestreo, debería explicarse la estrategia de muestreo en el informe de evaluación de la calidad.

Es preferible la inspección completa. En el caso en el que no sea factible, puede utilizarse la inspección por muestreo. El anexo E analiza en detalle los métodos de muestreo.

10.2.3 Evaluación indirecta

Un método de evaluación indirecto es un método de evaluación de la calidad de un conjunto de datos basado en conocimiento o experiencia externa sobre el producto de datos y puede ser subjetivo.

Ese conocimiento externo puede incluir información sobre el uso, linaje y propósito (véase la Norma ISO 19115-1) y otros informes sobre la calidad de los datos del conjunto de datos o de los datos usados para producirlo. La calidad de los datos puede estimarse, por ejemplo, a partir de información sobre la fuente, las herramientas y los métodos usados para capturar los datos, y evaluarse en contraste con procedimientos y especificaciones elaboradas para ese producto. La calidad de los datos evaluada de forma indirecta puede también basarse únicamente en la experiencia o determinarse a partir de los comentarios de los usuarios (por ejemplo, según se informa de acuerdo con la referencia [36]).

En algunos casos puede ser potencialmente engañoso o incluso imposible el informar sobre la calidad de los datos evaluada de forma indirecta con resultados cuantitativos. En tales casos la calidad de los datos puede describirse de forma textual usando un resultado descriptivo; véase el apartado 8.5.4.4.

10.3 Agregación y derivación

Se pueden generar resultados adicionales mediante la agregación o la derivación de resultados existentes sin llevar a cabo una nueva evaluación de la calidad de los datos.

La agregación combina resultados de evaluaciones de la calidad basadas en diferentes elementos de la calidad o diferentes ámbitos de la calidad.

También pueden derivarse resultados adicionales de resultados existentes, por ejemplo, cuando un resultado de conformidad se obtiene mediante la comparación de un resultado cuantitativo con un nivel de conformidad. Esto es útil si el resultado está expresado de forma diferente al nivel de conformidad, por ejemplo.

NOTA 1: La agregación puede usarse para agregar resultados de diferentes elementos de la calidad de los datos para describir la conformidad frente a unas especificaciones de producto.

NOTA 2: La agregación se describe más a fondo en el anexo G. En el apartado 11.2.1 y en el anexo D se describe cómo informar sobre la agregación.

NOTA 3: En el apartado 11.2.2 y en el anexo D se describe cómo informar sobre la derivación.

EJEMPLO: Si el resultado se expresa con una probabilidad del 95% y el nivel de conformidad se expresa con una probabilidad del 99%, el resultado podría recalcularse para tener la misma probabilidad que el nivel de conformidad.

11. INFORME DE LA CALIDAD DE LOS DATOS

11.1 Generalidades

Se debe informar sobre la calidad de los datos con metadatos y puede complementarse con un informe de evaluación de la calidad.

Requisito 5:	https://standards.iso/211.org/19157-1/1/req/content/dataQualityMetadata
	Se debe informar sobre la calidad de los datos con metadatos de acuerdo con el capítulo 8, el capítulo 11 y el anexo C de este documento, así como de acuerdo con la Norma ISO 19115-1.

Para proporcionar más detalles de los ofrecidos en los metadatos, e incluir detalles de la evaluación de la calidad de los datos, como los pasos específicos del proceso ejecutados en la evaluación, se puede crear un informe de evaluación de la calidad. La descripción de los metadatos incluye una referencia al informe de evaluación de la calidad cuando exista (véase la figura 13).

Requisito 6:	https://standards.iso211.org/19157-1/1/req/content/qualityEvaluationReport
	Un informe opcional de evaluación de la calidad no debe sustituir a los metadatos
Requisito 7:	https://standards.iso211.org/19157-1/1/rec/optionalContent/reportReference
	Si existe un informe de evaluación de la calidad para los metadatos de la calidad presentados en un informe, se debería proporcionar una referencia a este informe.

NOTA 1: Véase también el apartado B.4.3.2 para más información sobre cómo informar sobre la calidad de los datos y sobre el rol complementario entre los metadatos y el informe de evaluación de la calidad.

NOTA 2: Véase el capítulo E.4 para ejemplos sobre cómo informar sobre la calidad de los datos.

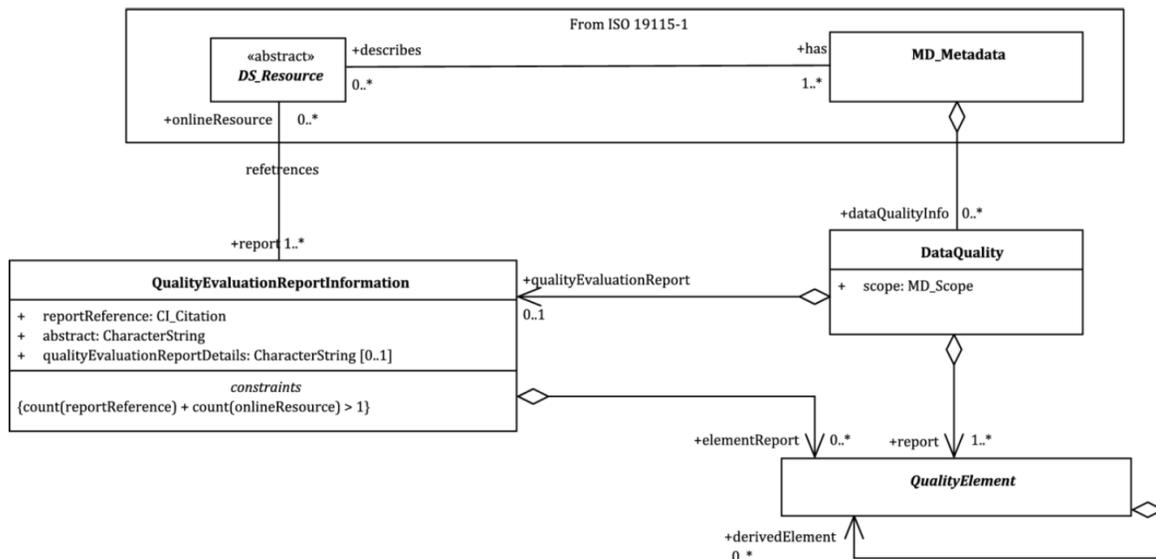


Figura 13 – Informe sobre la calidad de los datos

11.2 Casos particulares

11.2.1 Informe sobre la agregación (resultados agregados)

Un informe de evaluación de la calidad puede incluir información sobre los resultados agregados.

<p>Recomendación 8:</p>	<p>https://standards.isotc211.org/19157/-/1/1/rec/optionalContent/aggregatedResult</p>
	<p>Cuando se agrega el resultado, debería proporcionarse un informe de evaluación de la calidad para completar la información proporcionada de los metadatos.</p> <p>Este informe incluye información completamente detallada del resultado original [con la(s) medida(s) de evaluación], resultado agregado y método de agregación.</p>
<p>Recomendación 9:</p>	<p>https://standards.isotc211.org/19157/-/1/1/rec/optionalContent/sameQualityElement</p>
	<p>Cuando varios resultados de la calidad de un mismo elemento de la calidad se agregan en un resultado individual de ese elemento, el resultado debería consignarse en los metadatos como un resultado de ese elemento de la calidad de los datos. Véanse los apartados D.4.1.2 y D.4.1.3</p>
<p>Recomendación 10:</p>	<p>https://standards.isotc211.org/19157/-/1/1/rec/optionalContent/differentQualityElement</p>
	<p>Cuando varios resultados de la calidad de diferentes elementos de calidad se agregan en un resultado individual, este debería consignarse en los metadatos como un resultado de la evaluación de la conformidad del conjunto de datos con las especificaciones de producto o con los requerimientos del usuario. Véase el apartado D.4.1.4. para un ejemplo</p>

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Requisito 7:	https://standards.iso211.org/19157/-1/1/req/content/aggregatedResult
	<p>Cuando se consigna un resultado agregado en los metadatos, debe proporcionarse una referencia a los resultados de la evaluación de la calidad originales de los elementos individuales.</p>

11.2.2 Informe sobre la derivación (resultados derivados)

Cuando en los metadatos únicamente se informa de los resultados derivados, debería también generarse un informe de evaluación de la calidad para proporcionar los resultados de la calidad originales de los cuales se ha determinado el resultado derivado. En este caso los metadatos deberían proporcionar la referencia al informe de evaluación de la calidad y al resultado original de la calidad.

Recomendación 11:	https://standards.iso211.org/19157/-1/1/rec/optionalContent/qualityEvaluationReport
	<p>Para un resultado derivado se debería generar un informe de evaluación de la calidad que detalle la fuente de la derivación (por ejemplo, resultados de la calidad, método de derivación).</p>
Recomendación 12:	https://standards.iso211.org/19157/-1/1/rec/optionalContent/derivedResultMetadata
	<p>Los metadatos de un resultado derivado deberían proporcionar una referencia al informe de evaluación de la calidad y al resultado original de la calidad de los datos.</p>

Un resultado de conformidad suele derivar de un resultado cuantitativo. Si únicamente se informa el resultado de conformidad en los metadatos, los resultados cuantitativos deberían proporcionarse en un informe de evaluación de la calidad.

11.2.3 Referencia al resultado de la calidad original

Permiso 3:	https://standards.iso211.org/19157-1/1/per/permittedContent/sourceReference
	<p>Cuando se informa de los resultados agregados o derivados en los metadatos, puede proporcionarse la referencia al resultado original de la calidad de los datos usando dos atributos:</p> <ul style="list-style-type: none"> - El atributo <i>derivedElement</i> referencia un elemento de la calidad [y sus resultados] descrito en los metadatos. - El atributo <i>qualityEvaluationReportDetails</i> hace referencia a la parte del informe de evaluación de la calidad donde se describen los resultados originales.

11.2.4 Principio de jerarquía

Este documento reconoce el principio de nivel jerárquico.

La calidad de los datos especificada para niveles superiores (por ejemplo, una serie) es aplicable a un nivel inferior (por ejemplo, un conjunto de datos); véase la tabla 8.

Tabla 8 – Niveles jerárquicos

Nivel superior ↑	Serie	
	Conjunto de datos	
↓ Nivel inferior	Subconjunto	
	Tipo de objeto geográfico	Tipo de atributo
	Instancia de objeto geográfico	Instancia de atributo

NOTA: Puede informarse acerca de la calidad de una instancia de objeto geográfico, de atributo de objeto geográfico o de una asociación entre objetos geográficos como un atributo de la instancia tal y como se define en la Norma ISO 19109.

Recomendación 13:	https://standards.iso211.org/19157-1/1/rec/optionalContent/hierarchy
	Si la calidad de los datos difiere entre los niveles superior e inferior, entonces debería proporcionarse información complementaria para el nivel inferior.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

12. REQUISITO PARA LA CODIFICACIÓN XML

El intercambio de especificaciones de calidad de los datos en formato XML requiere de reglas de codificación. Este capítulo especifica esas reglas y un esquema XML. La clase de requisitos para la codificación XML se documenta en la tabla 3.

Requisito 8:	https://standards.iso211.org/19157/-1/1/req/xml
	Un documento XML de calidad de los datos debe ser conforme al esquema XML https://schemas.iso211.org/19157/-1/dqc/1.0.0/dqc.xsd .

El esquema XML sigue las reglas establecidas en la Especificación Técnica ISO/TS 19139-1 para la traducción de modelos UML al esquema XML.

El esquema XML también utiliza patrones para desvincular espacios de nombre XML descritos en el capítulo 8 de la Especificación Técnica ISO/TS 19115-3:2016.

Las definiciones del esquema XML pertenecen al siguiente espacio de nombres: <https://schemas.iso211.org/19157/-1/dqc/1.0>. Este espacio de nombres se abrevia dqc.

Este esquema XML aplica todas las clases UML definidas en este documento e importa todas las clases pertinentes de otras normas internacionales.

En el anexo H pueden encontrarse más detalles en relación con la codificación XML.

Anexo A
(Normativo)
Conjunto de pruebas abstractas

A.1 Contenido de unas especificaciones de producto de datos

Prueba de conformidad	https://standards.iso/211.org/19157/-1/1/conf/content/allContent
Referencia	Todas las declaraciones normativas en la clase de requisito. https://standards.iso/211.org/19157/-1/1/req/content (véase la tabla 4 en el capítulo 6)
Propósito de la prueba	Se verifica que el contenido de una definición de la calidad de los datos es conforme al modelo UML y los requisitos adicionales relacionados con los elementos del modelo UML.
Método de la prueba	Se identifica el objeto individual, sus atributos y sus relaciones con otros objetos. Se verifica que cada uno de estos elementos es conforme a la multiplicidad y tipo de datos expresado en el modelo UML.
Tipo de prueba	Básico

A.2 Codificación XML

Prueba de conformidad	https://standards.iso/211.org/19157/-1/1/conf/xml/xmlEncoding
Referencia	Todas las declaraciones normativas en la clase de requisito. https://standards.iso/211.org/19157/-1/1/req/xml (véase el capítulo 12 y el anexo H)
Propósito de la prueba	Se verifica la sintaxis de una definición de la calidad de los datos del documento XML.
Método de la prueba	Se valida el documento XML usando el esquema XML

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

	https://schemas.isotc211.org/19157-1/dqc/1.0.0/dqc.xsd
--	-----------------------------------------------------------------------------------------------------------------------------

CONSULTA PÚBLICA

**ANEXO B
(Normativo)
Conceptos sobre la calidad de los datos y su uso**

B.1 Marco de trabajo de los conceptos sobre la calidad de los datos

Un conjunto de datos puede estar producido para una aplicación específica o para un conjunto de aplicaciones supuestas. La calidad de un conjunto de datos solo puede evaluarse mediante el conocimiento de sus elementos de calidad de los datos y, en algunos casos, indirectamente mediante la información no cuantitativa de la calidad de los datos evalúan la diferencia entre el conjunto de datos y el universo de discurso (es decir, el conjunto de datos perfecto que se corresponde con las especificaciones de producto). La información no cuantitativa de la calidad proporciona información general de la que puede derivarse conocimiento relacionado con la calidad.

Los conceptos sobre la calidad de los datos proporcionan un importante marco de trabajo para los productores, así como para los usuarios. Al productor se le proporcionan los medios para validar en qué medida un conjunto de datos refleja su universo de discurso tal y como se en las especificaciones de producto. Los usuarios pueden evaluar la calidad de un conjunto de datos para determinar si puede satisfacer los requerimientos de la aplicación (véase la figura B.1).

Es importante señalar que los resultados de la calidad consignados son válidos frente a las especificaciones de producto o los requerimientos del usuario utilizados. Debería tenerse cuidado al comparar los resultados de la calidad cuando difiera el universo de discurso (es decir, las especificaciones o los requerimiento del usuario). Si estos cambiasen, la evaluación de la calidad debería repetirse respecto de las nuevas especificaciones o requisitos.

EJEMPLO: Ejemplos típicos de ello son los relacionados con la transformación de modelos en Infraestructuras de Datos Espaciales o en generalización. Por ejemplo, si cambia la geometría de un tipo de objeto geográfico, también cambian los resultados de exactitud posicional.

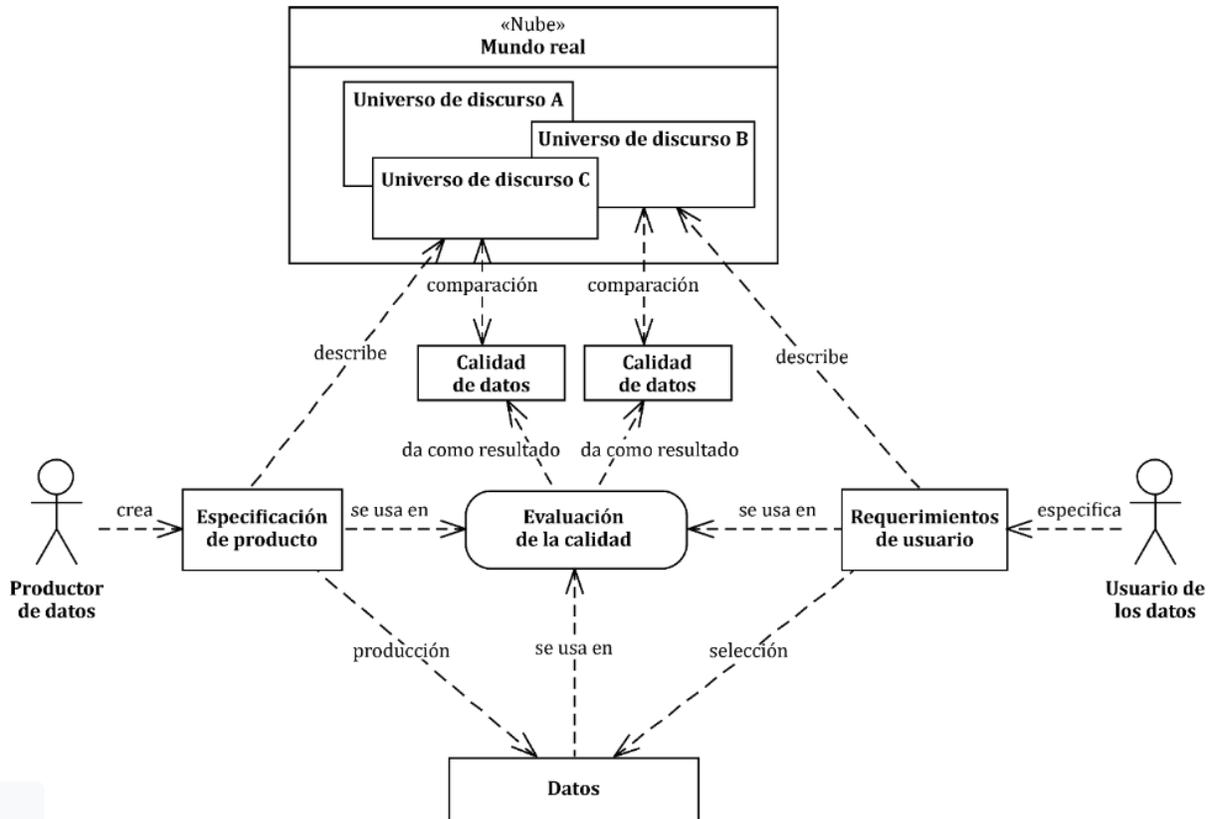


Figura B.1 – Marco de trabajo de los conceptos sobre la calidad de los datos

B.2 Estructura de los conjuntos de datos y componentes para la descripción de la calidad

Un conjunto de datos puede pertenecer a una serie, lo que significa que todos los conjuntos de datos de la serie se basan en las mismas especificaciones de producto de datos. La calidad de todos los conjuntos de la serie puede ser la misma.

Se puede considerar que un conjunto de datos contiene un número elevado, pero finito, de subconjuntos de datos. A menudo los subconjuntos de datos que compartan características comunes tienen calidades similares, como la pertenencia al mismo tipo de objeto geográfico, atributo de objeto geográfico o relación entre objetos geográficos, o que compartan criterios de agrupación o de extensión geográfica o temporal. Un subconjunto de datos puede ser tan reducido como una única instancia de objeto geográfico, un valor de atributo o un ocurrencia de relación entre objetos geográficos y, teóricamente, los conceptos usados para la calidad permiten que cada instancia, valor de atributo y ocurrencia de relación del conjunto de datos tenga su propia descripción de la calidad. La calidad de subconjuntos de un conjunto de datos no puede asumirse como idéntica a la calidad de otras partes del conjunto de datos al que pertenecen. Los conceptos utilizados sobre la calidad de los datos permiten informar sobre la calidad de un conjunto de datos y, adicionalmente, sobre la diferente calidad de subconjuntos de este conjunto de datos, identificando esas agrupaciones mediante la determinación del

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

ámbito de la calidad. La información suministrada sobre la calidad para diversos ámbitos menores que el conjunto de datos completo para el cual se informa sobre la calidad proporciona una visión más completa y detallada de la calidad que la calidad global del conjunto total de datos.

Para un productor de datos, unas especificaciones de producto describen el universo de discurso y contienen las reglas para construir el conjunto de datos. Para un usuario de los datos, los requerimientos de usuario describen un universo de discurso, que puede coincidir, o no, con el universo de discurso del conjunto de datos. La calidad de un conjunto de datos es la medida en que dicho conjunto representa un universo de discurso. Por lo tanto la calidad de un mismo conjunto de datos puede diferir dependiendo del universo de discurso frente al cual está siendo evaluado.

La calidad de un conjunto de datos se describe mediante los elementos de la calidad de los datos y sus descriptores. El propósito, uso y linaje también pueden proporcionar alguna información relacionada con la calidad. La metacalidad proporciona información sobre la calidad de la evaluación de la calidad.

Los elementos de la calidad permiten evaluar el grado de adecuación de un conjunto de datos a los criterios establecidos en sus especificaciones de producto o requerimientos de usuario. Los elementos pueden evaluarse de diferentes formas y en diferentes etapas del ciclo de vida de un conjunto de datos. No todos los elementos de la calidad son aplicables a todos los tipos de conjuntos de datos. Algunos elementos son aplicables a conjuntos de datos extensos, mientras otros son más adecuados para subconjuntos de conjuntos de datos mayores. Algunos elementos de la calidad son aplicables a instancias individuales de datos, así como a un número elevado, mientras que algunos solo son aplicables a múltiples instancias.

Este documento fundamentalmente identifica los elementos de la calidad de los datos como un medio para identificar e informar sobre categorías independientes de la información sobre la calidad. Adicionalmente reconoce que con frecuencia los elementos de la calidad están interrelacionados. Por ejemplo, un error en coordenadas puede generar al menos dos tipos de errores: un error posicional y un error topológico, véase anexo G. Es responsabilidad del evaluador de la calidad el significado de los elementos de la calidad y la forma en que se manejan.

Todos los elementos de la calidad pueden usarse para evaluar la conformidad del conjunto de datos frente a las especificaciones de producto o los requerimientos definidos por el usuario. La evaluación puede basarse en una aplicación específica o en unos requerimientos de usuario que no pueden describirse utilizando los elementos de la calidad descritos anteriormente. En este caso, pueden usarse componentes definidos en otros documentos de la serie (por ejemplo, LI_Lineage de la Norma ISO 19115-1) o pueden definirse nuevos elementos según los requisitos establecidos en el apartado 8.4 para nuevos elementos de la calidad. Asimismo, pueden referenciarse elementos normalizados existentes definidos en otro lugar.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Los productores de datos pueden mostrar cómo un conjunto de datos es adecuado para varios usos identificados. Esto se puede hacer como texto libre en una identificación del conjunto de datos (en el *abstract* de la clase MD_Identification de la Norma ISO 19115-1) o, cuando se necesita una explicación más detallada y estructurada, con la clase MD_Usage de la Norma ISO 19115-1.

Los requisitos para la adecuación al uso de un conjunto de datos para un uso específico definido por el productor pueden proporcionarse en las especificaciones de producto definiendo niveles de conformidad de la calidad para los valores de los elementos de la calidad pertinentes.

EJEMPLO: Los registros de la propiedad y el correspondiente mapa digital de propiedades son adecuados para la constitución de propiedades en el sistema de referencia local dentro de zonas urbanizadas.

Exactitud posicional local de la linde de bienes inmuebles:	0,05 m
Exactitud posicional externa de la linde de bienes inmuebles:	0,2 m
Omisión de unidades de bienes conjuntos:	< 1%

B.3 Cuándo utilizar los procedimientos de evaluación de la calidad

Los procedimientos de evaluación de la calidad pueden usarse en diferentes fases del ciclo de vida de un producto. Las etapas del ciclo de vida de un producto, durante las que puede aplicarse la evaluación de la calidad, son las siguientes:

- Desarrollo de las especificaciones de producto o de los requerimientos de usuario: Cuando se desarrollan las especificaciones de un producto o se definen los requerimientos de usuario, los procedimientos de evaluación de la calidad pueden utilizarse para facilitar el establecimiento de los niveles de conformidad de la calidad que debería cumplir el producto final. Unas especificaciones de producto de datos, o unos requerimientos de usuario, pueden incluir niveles de conformidad para la calidad de los datos y procedimientos de evaluación de la calidad a aplicar durante la producción y actualización.
- Control de calidad durante la creación del conjunto de datos: durante la etapa de producción, como parte del proceso de control de calidad, el producto puede aplicar procedimientos de evaluación de la calidad establecidos explícitamente, o no, en las especificaciones de producto de datos. Cuando se utilizan para el control de calidad en la producción, puede informarse en los metadatos con una descripción de los procedimientos aplicados de evaluación de la calidad, en forma de linaje, incluyendo, pero no necesariamente limitado a estos, los procedimientos de evaluación de la calidad aplicados, los niveles establecidos de conformidad de la calidad y los resultados.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Inspección para determinar la conformidad con unas especificaciones de producto: Una vez finalizada la producción, puede utilizarse un proceso de evaluación de la calidad para producir e informar sobre los resultados de la calidad de los datos. Estos resultados pueden utilizarse para determinar si un conjunto de datos es conforme con sus especificaciones de producto. Si el conjunto de datos supera la inspección (compuesta por un conjunto de procedimientos de evaluación de la calidad), se considera listo para su uso. Debería informarse sobre los resultados de la inspección de acuerdo con el capítulo 11. Véase también el ejemplo del anexo D que describe la evaluación e informe sobre la calidad de los datos. A consecuencia de una inspección se producirá la aceptación o rechazo del conjunto de datos. Si se rechaza, tras la corrección pertinente se requerirá una nueva inspección antes de que el producto pueda considerarse conforme con las especificaciones de producto.
- Evaluación de la conformidad de un conjunto de datos respecto a los requerimientos de usuario: Pueden usarse procedimientos de evaluación de la calidad para determinar si un conjunto de datos cumple los niveles de conformidad de la calidad especificados en los requerimientos de usuario. En los análisis de conformidad pueden usarse tanto métodos directos como indirectos.
- Control de calidad durante la actualización de los datos: Los procedimientos de evaluación de la calidad se aplican a las operaciones de actualización de los datos, tanto a los ítems usados para la actualización como a aquellos usados para determinar la calidad del conjunto de datos tras la actualización.

B.4 Informas sobre la calidad

B.4.1 Por qué informar sobre la calidad

La necesidad de informar sobre la calidad existe por numerosas razones, incluyendo las siguientes.

- Para ayudar a encontrar el conjunto de datos y fomentar su uso;
- Para demostrar la conformidad con una especificaciones de producto o unos requerimientos de usuario;
- Como parte de las iniciativas de gestión de los proveedores;
- Para permitir posteriores dictámenes sobre la calidad de la información derivada del conjunto de datos;
- Para permitir la toma racional (óptima) de decisiones cuando se sabe que todos los datos contienen defectos.

B.4.2 Cuándo informar sobre la calidad

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Continuamente se están creando, actualizando y fusionando conjuntos de datos, por lo que su calidad o una de sus componentes pueden variar. La calidad de un conjunto de datos puede verse afectada bajo tres circunstancias:

- Cuando se elimina, modifica o añade cualquier cantidad de datos al conjunto,
- Cuando se modifican las especificaciones de producto de datos o si se identifican nuevos requerimientos de usuario,
- Cuando cambia el mundo real

La primera circunstancia, una modificación del conjunto de datos, puede ocurrir con bastante frecuencia. Muchos conjuntos de datos no son estáticos. El intercambio de información aumenta, así como el empleo de unos mismos datos para múltiples propósitos y la consecuente actualización y refinamiento para alcanzar esos propósitos. Dado que es probable que el informe de la calidad de un conjunto de datos tenga que cambiar por las modificaciones ocurridas en el conjunto, la calidad debería evaluarse y actualizarse cuando sea necesario al producirse cambios.

Al crearse un conjunto de datos, deberían conocerse completamente todos los elementos de la calidad aplicables. Inicialmente el productor solo puede informar sobre el uso (asumiendo que le productor realmente utiliza dicho conjunto de datos). Se depende de los usuarios de los datos para que informen sobre usos distintos a los inicialmente previstos, por lo que se pueden realizar continuas actualizaciones de este elemento general de la calidad de los datos para contemplar los usos no previstos que están teniendo lugar.

La segunda circunstancia, una modificación de las especificaciones de producto de datos, es más probable que ocurra antes de la construcción inicial del conjunto de datos y previamente a la publicación de la información sobre la calidad. Sin embargo, es concebible que, conforme se utilice un conjunto de datos, se actualicen sus especificaciones para que las futuras modificaciones satisfagan mejor las necesidades reales. Al modificar las especificaciones de producto, la calidad del conjunto de datos actual también cambia. La información sobre la calidad de un conjunto de datos siempre debería contemplar el conjunto de datos y las especificaciones actuales de producto de datos.

La tercera circunstancia, un cambio en el mundo real, ocurre de forma continua. Los cambios pueden estar causados por fenómenos naturales, como es el caso de los movimientos de la corteza terrestre o la erosión, pero generalmente son un resultado de la actividad humana. A menudo los cambios son muy rápidos y drásticos. Por este motivo, al determinar la calidad de un conjunto de datos, la fecha de la toma de los datos es tan importante como la fecha de la evaluación de la calidad. En algunos casos, si se conoce, también es de interés el índice de cambio. La frecuencia de actualización del conjunto de datos también puede ser de interés en algunos casos. No obstante, este documento reconoce que potencialmente podrá no ser posible crear un nuevo informe de la calidad de los datos cada vez que cambie el mundo real.

Este documento comprende un borrador de Proyecto de Norma Técnica diseñado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

CONSULTA PÚBLICA

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

B.4.3 Cómo informar sobre la calidad con metadatos y con un informe de evaluación de la calidad

B.4.3.1 Generalidades

Se puede informar sobre la calidad con metadatos y mediante un informe de evaluación de la calidad.

- Los metadatos pretenden proporcionar información breve, sintética y estructurada de forma general para permitir la interoperabilidad de los metadatos y el uso de servicios web.
- El informe de evaluación de la calidad puede usarse para proporcionar información completa y detallada sobre la evaluación de la calidad de los datos. El informe tiene que referenciarse desde los metadatos como se explica en el capítulo 11.

Por ejemplo, en el caso de la agregación de diferentes resultados de la calidad, el informe proporcionará información completa de los resultados originales (con los procedimientos de evaluación y medidas aplicadas), del resultado agregado y del método de agregación, mientras que los metadatos potencialmente solo pueden describir el resultado agregado con una referencia a los resultados originales que se describen en el informe.

B.4.3.2 Reporte de la información sobre la calidad como metadatos

La clase MD_Metadata, definida en la Norma ISO 19115-1, agrega ninguna o varias unidades de la calidad de los datos (instancias de la clase DataQuality, tal y como se especifica en este documento), véase la figura B.2.

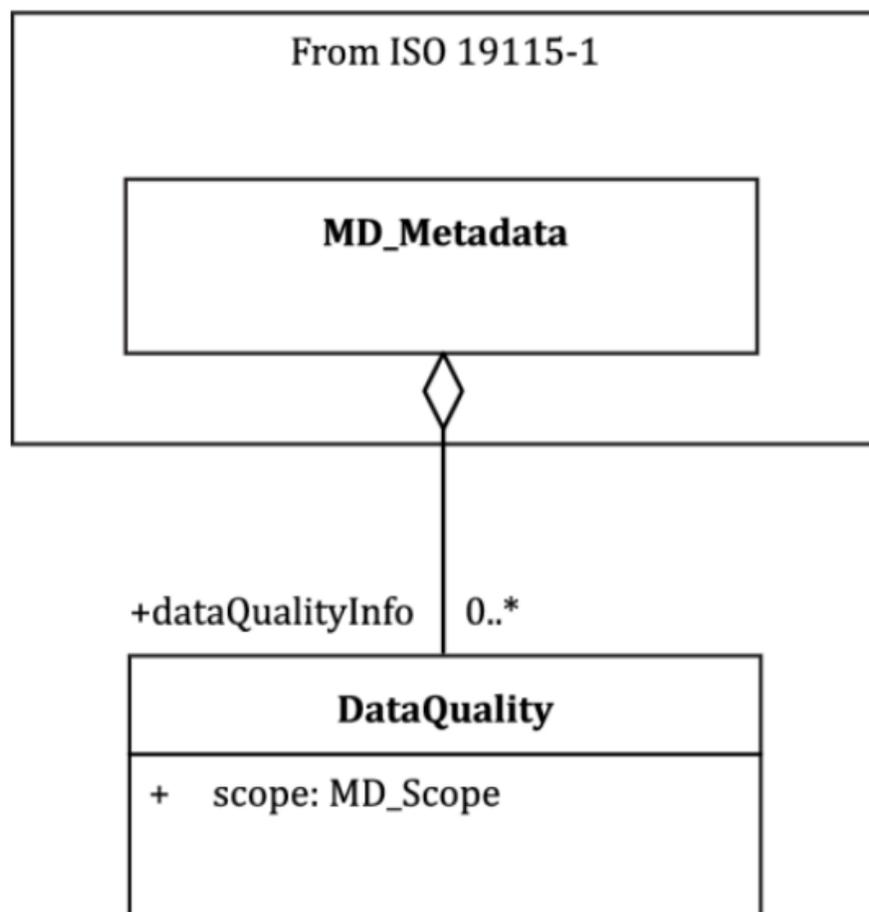


Figura B.2 – Información sobre la calidad de los datos

B.4.3.3 Reporte de información sobre la calidad en un informe de evaluación de la calidad

La normalización de la terminología (por ejemplo, los elementos de calidad y de los datos) y de la estructura de la información subyacente de la calidad de los datos será beneficiosa para los usuarios familiarizados con la norma y facilitará una mejor comprensión y comparación. Además, puede ser de valor para los usuarios incluir en el informe una declaración de la conformidad con la norma.

Un informe de evaluación de la calidad debería contener un ámbito para identificar con facilidad la extensión cubierta por el informe en relación con los datos que son evaluados.

Cada informe debería contener la información suficiente para describir de forma significativa los aspectos relevantes de la calidad de los datos y los resultados. Puede consistir en referencias a documentación adicional, tal como las especificaciones de producto de datos o un registro de medidas.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

La estructura completa del informe de evaluación de la calidad no se ha normalizado intencionalmente, por lo que cada organización particular puede adaptarla para sus propias necesidades, prácticas y procedimientos de evaluación. El informe puede estar en texto libre. No obstante, puede ser importante la cantidad de información sobre la calidad. Por tanto, es importante presentar el informe de forma sucinta, fácil de comprender y de recuperar. Por ejemplo, un informe puede seguir la estructura que se describe en este documento. En el anexo D se proporciona un ejemplo de informe de evaluación de la calidad.

CONSULTA PÚBLICA

**ANEXO C
(Normativo)**

Diccionario de datos para la calidad de los datos

C.1 Visión general del diccionario de datos

C.1.1 Introducción

Este diccionario de datos describe las características del modelo de la calidad de los datos definidas en los capítulos 8, 9, 10 y 11. El diccionario se especifica en tablas, con columnas y filas en una jerarquía para establecer relaciones y una organización para la información.

Las filas sombreadas de la tabla representan las clases. Las filas no sombreadas de la tabla representan atributos de clase y asociaciones. Las clases y atributos de clases dentro de las tablas del diccionario de datos están definidas por seis columnas que se describen en los apartados C.1.2 a C.1.7.

C.1.2 Nombre/nombre del rol

Es una etiqueta asignada a la clase o atributo de clase. Los nombres de clase comienzan con una letra en mayúscula. No aparecen espacios en un nombre de clase. En cambio, varias palabras se concatenan y cada nueva subpalabra comienza con una letra mayúscula (ejemplo: XnnnYmmm). Los nombres de clase son únicos en todo el diccionario de datos de este documento. Los nombres de atributo de clase son únicos dentro de una clase, pero no en todo el diccionario de datos de este documentos. Los nombres de atributo de clase de hacen únicos, dentro de una aplicación, por la combinación del nombre de clase. Los nombres y los nombres de rol pueden estar en un idioma distinto del que se utiliza en este documento.

C.3.1 Definición

Esta es la descripción de la clase o del atributo de clase.

C.1.4 Obligación/condición

C.1.4.1 Generalidades

Este es un descriptor que indica si una clase o un atributo de clase debe siempre documentarse en el conjunto de datos o solo en algunas ocasiones [es decir, contiene valor(es)]. Este campo puede tomar los siguientes valores: O (obligatorio), C (condicional) u Op (opcional).

C.1.4.2 Obligatorio (O)

Este descriptor especifica que la clase o atributo de clase debe estar documentada siempre.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

C.1.4.3 Condicional (C)

Este descriptor especifica una condición gestionada electrónicamente bajo la cual al menos una clase, un atributo de clase o una asociación es obligatoria. “Condicional” se usa en una de la tres posibilidades siguientes.

- Para expresar una elección entre dos o más opciones. Al menos una opción es obligatoria y debe documentarse.
- Para documentar una clase, un atributo de clase , un atributo de clase o una asociación si otra clase ha sido documentada.
- Para documentar un atributo de clase o asociación si un valor específico para otro atributo de clase ha sido documentado. Para facilitar la interpretación, el valor específico se expresa con un texto plano. Sin embargo, el código se debe utilizar para verificar la condición en una interfaz de usuario digital.

Si la respuesta a la condición es positiva, entonces la clase, el atributo de clase o asociación debe ser obligatorio.

C.1.4.4 Opcional (O)

Este descriptor especifica que la clase, el atributo de la clase o la asociación pueden documentarse o no. Las clases opcionales o los atributos de clase opcionales se han definido para proporcionar una orientación para los que quieren documentar completamente sus datos. (El uso de este conjunto común de elementos definidos ayudará a promocionar la interoperabilidad entre usuarios y productores de datos geográficos de todo el mundo). Si no se usa una entidad opcional, los atributos de clase contenido en esa clase (incluyendo atributos obligatorios) no se usarán. Las clases opcionales pueden tener atributos de clase obligatorios; esos atributos de clase solo son obligatorios si es usa la clase opcional.

C.1.5 Ocurrencia máxima

Este descriptor especifica el número máximo de instancias que la clase, atributo de clase o asociación puede tener. Las ocurrencias singulares son mostradas como “1” y si se repiten las ocurrencias se representa mediante “N”. Se permiten ocurrencias de número fijo diferente a “1” y se representarán con su número correspondiente (es decir, “2”, “3”, etcétera).

C.1.6 Tipo de datos

Este descriptor especifica un conjunto de valores bien diferenciados para representar los atributos de clase, por ejemplo, *integer* (entero), *real* (real), *string* (cadena de caracteres), *DataTime* (fecha y hora) y *Boolean* (Booleano). El atributo Tipo de datos se usa también para definir clases, estereotipos y asociaciones de clases.

NOTA: Los tipos de datos se definen en la Especificación Técnica ISO 19103:2015, apartado 6.5.2. Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

C.1.7 Dominio

Para una clase (fila sombreada), el dominio indica los números de línea cubiertos por los atributos de clase y asociaciones para esa clase.

Para un atributo de clase o asociación, el dominio especifica los valores permitidos. El uso de texto libre indica que el contenido de un campo puede comportar información textual sin restricciones; “dominio sin especificar” significa que puede utilizarse cualquier conjunto de caracteres alfanuméricos.

NOTA: Los dominios se definen en la Norma ISO 19115-1:2014, apartado B.1.7.

C.2 Diccionario de datos del paquete de calidad de los datos

C.2.1 Calidad de los datos

C.2.1.1 Generalidades

El modelo UML global para el paquete de calidad de los datos completo se muestra en la figura 2.

En la tabla C.1 se describen las propiedades del modelo UML que se muestra en la figura 3 y la figura 13.

Tabla C.1 – Calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
1	DataQuality	Grado con el que un conjunto de características inherentes de los datos cumple unos requisitos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase agregada (MD_Metadata)	Líneas 2-4
2	Scope	Los datos específicos a los que se refiere la información de la calidad de los datos	0	1	Clase	MD_Scope << DataType >> (ISO 19115-1)
3	<i>Role name:</i> report	Una referencia al informe (metadatos o informe de evaluación de la calidad) asociado con el elemento de la calidad.	0	N	Asociación	QualityElement (Clase abstracta) (C.2.1.2)
4	<i>Role name:</i> QualityEvaluationReport	Una referencia al informe (metadatos o informe de evaluación de la calidad) asociado con el elemento de la calidad.	Op	1	Asociación	QualityEvaluationReportInformation (C.2.1.6)

C.2.1.2 Elemento de calidad de los datos

En la tabla C.2 se describe el modelo UML que se muestra en la figura 2, figura 4, la figura 5, la figura 10 y la figura 13.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla C.2 – Elemento de calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
5	<i>QualityElement</i>	Aspecto de la información cuantitativa de la calidad.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase agregada (DataQuality) (Clase Abstracta)	Líneas 6-9
6	<i>Role name: evaluation Method</i>	Información sobre la evaluación.	0	N	Asociación	EvaluationMethod (C.2.1.4)
7	<i>Role name: result</i>	Valor (o conjunto de valores) obtenidos de aplicar una unidad de la calidad de los datos o el resultado de evaluar el valor obtenido (o conjunto de valores) frente a un nivel de conformidad de la calidad especificado como aceptable	0	N	Asociación	QualityResult (Clase Abstracta) (C.2.1.5)
8	<i>Role name: derived Element</i>	En caso de agregación o derivación indica el elemento original.	Op	N	Asociación	QualityElement (Clase Abstracta) (C.2.1.2)

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
9	<i>Completeness</i>	Presencia y ausencia de objetos geográficos, sus atributos y sus relaciones.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (QualityElement (Clase Abstracta))	Líneas 6-9
10	<i>Commission</i>	Exceso de datos presentes en el conjunto de datos según se describe en el ámbito.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Completeness)	Líneas 6-9
11	<i>Omission</i>	Ausencia de datos presentes en el conjunto de datos según se describe en ámbito.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Completeness)	Líneas 6-9
12	<i>Logical-Consistency</i>	Grado de adherencia a reglas lógicas de la estructura de datos, de los atributos y de las relaciones (la estructura de los datos puede ser conceptual, lógica o física).	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Quality Element) (Clase Abstracta)	Líneas 6-9
13	<i>Conceptual-Consistency</i>	Adherencia a las reglas del esquema conceptual.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Logical Consistency)	Líneas 6-9

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

14	Domain-Consistency	Adherencia de los valores a los dominios de los valores.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Logical Consistency)	Líneas 6-9
15	Format-Consistency	Grado en que los datos se almacenan de acuerdo con la estructura física del conjunto de datos según se describe por el ámbito.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Logical Consistency)	Líneas 6-9
16	TopologicalConsistency	Corrección de las características topológicas del conjunto de datos codificadas explícitamente según se describe en el ámbito.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Logical Consistency)	Líneas 6-9
17	Positional-Accuracy	Exactitud de la posición de los objetos geográficos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Quality Element) (Clase Abstracta)	Líneas 6-9
18	AbsolutePositionalAccuracy	Proximidad de los valores consignados de las coordenadas a los valores aceptados como verdaderos en un sistema de referencia de coordenadas estandarizado.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Positional Accuracy)	Líneas 6-9

19	RelativePositionalAccuracy	Proximidad de las posiciones relativas de los objetos geográficos del ámbito a sus respectivas posiciones relativas aceptadas como verdaderas en un sistema de referencia de coordenadas local.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Positional Accuracy)	Líneas 6-9
20	GriddedDataPositionalAccuracy	Proximidad de los valores de posición de los datos en estructura de malla a los valores aceptados como verdaderos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Positional Accuracy)	Líneas 6-9
21	Thematic-Accuracy	Exactitud de los atributos cuantitativos, corrección de los atributos no cuantitativos y corrección de las clasificaciones de objetos geográficos y sus relaciones.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Quality Element) (Clase Abstracta)	Líneas 6-9
22	Thematic-ClassificationCorrectness	Comparación de las clases asignadas a los objetos geográficos o a sus atributos frente a un universo de discurso.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Thematic Accuracy)	Líneas 6-9
23	NonQuantitativeAttributeCorrectness	Medida que permite establecer si un atributo no cuantitativo es correcto o incorrecto.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Thematic Accuracy)	Líneas 6-9

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

24	QuantitativeAttribute Accuracy	Proximidad del valor de un atributo cuantitativo al valor verdadero o al aceptado como tal.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Thematic Accuracy)	Líneas 6-9
25	TemporalQuality	Exactitud de los atributos temporales y de las relaciones temporales de los objetos geográficos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Quality Element) (Clase Abstracta)	Líneas 6-9
26	AccuracyOfATimeMeasurement	Proximidad de las mediciones de tiempo consignadas a los valores aceptados como verdaderos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Temporal Quality)	Líneas 6-9
27	TemporalValidity	Validez de los datos con respecto al tiempo	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Temporal Quality)	Líneas 6-9
28	Metaquality	Información relativa a la fiabilidad de los resultados de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Quality Element) (Clase Abstracta)	Líneas 30 y 6-9

29	Role name: relatedQualityElement	Elemento de la calidad para el que se aplica la información de metacalidad.	0	1	Asociación	QualityElement (Clase Abstracta) (C.2.1.2)
30	Confidence	Fiabilidad de un resultado de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Metaquality)	Líneas 30 y 6-9
31	Representativity	Grado en que la muestra utilizada ha producido un resultado que es representativo de los datos pertenecientes al ámbito.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Metaquality)	Líneas 30 y 6-9
32	Homogeneity	Uniformidad esperada o comprobada de los resultados obtenidos para una evaluación de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Metaquality)	Líneas 30 y 6-9

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

C.1.2.3 Referencia de medida

En la tabla C.3 se describe el modelo UML que se muestra en la figura 6.

Tabla C.3 – Referencia de medida

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
33	Measure-Reference	Referencia a la medida utilizada.	Utiliza la obligación del objeto de referencia	Utiliza la máxima ocurrencia del objeto de referencia.	Clase agregada (Element)	Líneas 35-37
34	measureIdentification	Identificador de la medida, valor que identifica únicamente una medida dentro de un espacio de nombres.	Op	1	Clase	MD_Identifier << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.17.2)
35	nameOfMeasure	Nombre de la prueba aplicada a los datos.	C/si measure-Identification no se documenta.	N	Cadena de caracteres	Texto libre
36	measureDescription	Descripción de la medida.	C/si measure-Identification no se documenta.	1	Cadena de caracteres	Texto libre

C.2.1.4 Evaluación de la calidad de los datos

En la tabla C.4 se describe el modelo UML que se muestra en la figura 7, la figura 11 y la figura 12.

Tabla C.4 – Evaluación de la calidad de los datos

	Nombre/rol	Definición	Obligación/ condición	Máxima ocurrencia	Tipo de datos	Dominio
37	DQ_Evaluation-Method	Descripción del método de evaluación y del procedimiento aplicado.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase agregada (Quality Element)	Líneas 39-45
38	name	Nombre del método de evaluación.	Op	1	Cadena de caracteres	Texto libre
39	evaluationMethodDescription	Descripción del método de evaluación.	Op	1	Cadena de caracteres	Texto libre
40	evaluationMethodType	Tipo de método usado para evaluar la calidad de los datos	Op	1	Clase	Evaluation-MethodType Code << CodeList >> (C.3.2)
41	evaluationProcedure	Referencia a la información del procedimiento.	Op	1	Clase	CI_Citation << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.16)
42	dateTime	Fecha o rango de fechas en los que se ha aplicado una medida de calidad de los datos	Op	N	Clase	DateTime (véase la Norma ISO 19103:2015)
43	referenceDoc	Información sobre los documentos que se referencian en el desarrollo y aplicación de un método de evaluación de la calidad de los datos	Op	N	Clase	CI_Citation << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.16)

44	deductiveSource	Información sobre qué datos son usados como fuente en un método de evaluación deductivo	0	1	Cadena de caracteres	Texto libre
45	DataEvaluation	Método de evaluación de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (EvaluationMethod) (Clase Abstracta)	Líneas 39-45
46	FullInspection	La evaluación de la calidad comprueba cada ítem especificado por el ámbito de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Data-Evaluation)	Líneas 39-45
47	SampleBasedInspection	La evaluación de la calidad se realiza en subconjuntos de los datos geográficos definidos por el ámbito de la calidad de los datos	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Data-Evaluation)	Líneas 39-45 y 49-51
48	samplingScheme	Información sobre el tipo de esquema de muestreo y descripción del procedimiento de muestreo.	0	1	Cadena de caracteres	Texto libre

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

49	lotDescription	Información sobre cómo se definen los lotes.	0	1	Cadena de caracteres	Texto libre
50	samplingRatio	Información sobre cuántas muestras se extraen, en promedio, de cada lote de la población para su inspección.	0	1	Cadena de caracteres	Texto libre
51	Indirect-Evaluation	Evaluación de la calidad basada en conocimiento o experiencia externa sobre el producto de datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Data-Evaluation)	Líneas 39-45
52	AggregationDerivation	Método de agregación o de derivación.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Evaluation)	Líneas 39-45

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

C.2.1.5 Resultado de la calidad de los datos

En la tabla C.5 se describe el modelo UML que se muestra en la figura 8.

Tabla C.5 – Resultado de la calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
53	<i>QualityResult</i>	Generalización de clases más específicas de resultados.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase agregada (Element) (Clase Abstracta)	Líneas 55-56
54	resultScope	Ámbito del resultado.	Op	1	Clase	MD_Scope (ISO 19115-1:2014)
55	dateTime	Fecha en la que se generó el resultado.	Op	1	Clase	DateTime (véase la Norma ISO 19103:2015)
56	QuantitativeResult	Los valores o la información sobre el(los) valor(es) (o conjunto de valores) obtenidos de aplicar una medida de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Result)	Líneas 58-60 y 55-56
57	value	Valor o valores cuantitativos, cuyo contenido está determinado por el procedimiento de evaluación utilizado, de acuerdo con el valueType y valueStructure definidos para la medida	O	N	Clase	Record (véase la Norma ISO 19103:2015)
58	valueUnit	Unidad del valor para reportar un resultado de la calidad de los datos.	C/Si el valor se describe con un número	1	Clase	UnitOfMeasure (véase la Norma ISO 19103:2015)
59	valueRecordType	Tipo de valor para reportar un resultado de la calidad de los datos, que depende de la implementación.	Op	1	Clase	RecordType << Metaclass >> (véase la Norma ISO 19103:2015)
60	ConformanceResult	Información sobre el resultado de evaluar el valor obtenido (o conjunto de valores) frente a un nivel de conformidad de la calidad específico y aceptable.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (QualityResult)	Líneas 62-64 y 55-56
61	specification	Mención de las especificaciones del producto de datos o del requerimiento de usuario frente al que se evalúan los datos.	O	1	Clase	CI_Citation << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.16)
62	explanation	Explicación del significado de la conformidad para este resultado.	Op	1	Cadena de caracteres	Texto libre

Este documento comprende un borrador de Proyecto de Norma Técnica diseñado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

63	pass	Indicación del resultado de conformidad donde 0 = no apto y 1 = apto.	0	1	Booleano	1 = sí 0 = no
64	Descriptive-Result	Resultado descriptivo de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Result)	Líneas 66 y 55-56
65	statement	Expresión textual del resultado descriptivo.	0	1	Cadena de caracteres	Texto libre

66	Coverage-Result	Resultado de una medida de calidad de los datos que organiza los valores medidos como cobertura.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase especificada (Result)	Líneas 68-71 y 55-56
67	spatialRepresentationType	Método utilizado para representar espacialmente el resultado en cobertura.	M	1	Clase	MD_SpatialRepresentationTypeCode < <CodeList> > (ISO 19115-1:2014/ Amd 1:2018)
68	Role name: resultSpatialRepresentation	Proporciona la representación digital de medidas de calidad de los datos que componen el resultado en cobertura.	M	1	Asociación	MD_SpatialRepresentation < <Abstract> > (ISO 19115-1:2014/ Amd 1:2018)

69	Role name: resultContent	Proporciona la descripción del contenido del resultado en cobertura cuando la cobertura de calidad se incluye con el recurso que se describe, es decir, la definición semántica de las medidas de calidad de los datos.	C/si no se proporciona resultFormat	N	Asociación	MD_RangeDimension (ISO 19115-1:2014)
70	Role name: resultFormat	Proporciona información sobre el formato de los datos del resultado en cobertura.	C/si no se proporciona resultContent	1	Asociación	MD_Format (ISO 19115-1:2014)

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

C.2.1.6 Contenido del informe de evaluación de la calidad

En la tabla C.6 se describe el modelo UML que se muestra en la figura 13.

Tabla C.6 – Contenido del informe de evaluación de la calidad

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
71	QualityEvaluationReportInformation	Referencia a un informe de evaluación de la calidad externo.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Líneas 73-77
72	reportReference	Referencia al informe de evaluación de la calidad asociado.	C/si no se proporciona onlineResource	1	Clase	CI_Citation << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.16)
73	abstract	Resumen del informe de evaluación de la calidad asociado.	0	1	Cadena de caracteres	Texto libre
74	qualityEvaluationReportDetails	Referencia a los resultados originales en el informe de evaluación de la calidad.	Op	N	Cadena de caracteres	Texto libre
75	Role name: elementReport	Elementos de la calidad que abarca el informe de evaluación de la calidad.	Op	N	Asociación	QualityElement
76	Role name: onlineResource	Referencia a un recurso en línea.	C/si no se proporciona reportReference	N	Asociación	DS_Resource << Abstract >> (véase la Norma ISO 19115-1:2014)

C.2.2 Medida de la calidad de los datos

C.2.2.1 Generalidades

El modelo UML para la información sobre las medidas se muestra en la figura 10.

C.2.2.2 Medidas de calidad de los datos

En la tabla C.7 se describe el modelo UML que se muestra en la figura 10.

Tabla C.7 – Medidas de calidad de los datos

	Nombre/rol	Definición	Obligación/ condición	Máxima ocurrencia	Tipo de datos	Dominio
77	QualityMeasure	Medida de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Líneas 79-90
78	measureIdentifier	Valor que identifica únicamente una medida dentro de un espacio de nombres.	0	1	Clase	MD_Identifier << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.17.2)
79	name	Nombre de la medida de la calidad de los datos aplicada a los datos.	0	1	Cadena de caracteres	Texto libre
80	sobrenombre	Otro nombre reconocido, una abreviatura o un nombre corto para la misma medida de la calidad de los datos.	Op	N	Cadena de caracteres	Texto libre
81	elementName	Nombre del elemento de la calidad de los datos para el que se informa de la calidad.	0	N	Clase	TypeName << interface >> (véase la Norma ISO 19103:2015)
82	definition	Definición del concepto fundamental de la medida de la calidad de los datos.	0	1	Cadena de caracteres	Texto libre
83	description	Descripción de la medida de la calidad de los datos, incluyendo todas las fórmulas e ilustraciones necesarias para obtener el resultado de aplicar la medida.	0	N	Clase	MeasureDescription << Datatype >> (C.2.2.5)
84	valueType	Tipo de valor para reportar un resultado de la calidad de los datos (debe ser uno de los tipos de datos definidos en la Norma ISO 19103:2015)	0	1	Clase	TypeName << interface >> (véase la Norma ISO 19103:2015)
85	valueStructure	Estructura para reportar un resultado de la calidad de los datos complejo.	Op	1	Clase	ValueStructure << CodeList >>(C.3.3)

86	example	Ilustración del uso de una medida de la calidad de los datos.	Op	N	Clase	MeasureDescription (C.2.2.5)
87	<i>Role name:</i> basicMeasure	Nombre de la medida básica de calidad de los datos de la que se deriva la medida de la calidad de los datos.	C/ Si se deriva de una medida básica.	1	Asociación	BasicMeasure (C.2.2.3)
88	<i>Role name:</i> sourceReference	Referencia a la fuente de un ítem que se ha adoptado de una fuente externa.	C / Si existe una fuente externa.	N	Asociación	SourceReference (C.2.2.6)
89	<i>Role name:</i> parameter	Variable auxiliar utilizada por la medida de la calidad de los datos, incluyendo su nombre, definición y opcionalmente su descripción.	C / Si se requiere.	N	Asociación	MeasureParameter (C.2.2.4)

C.2.2.3 Medida básica de calidad de los datos

En la tabla C.8 se describe el modelo UML que se muestra en la figura 10.

Tabla C.8 – Medida básica de calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
90	BasicMeasure	Medida básica de calidad de los datos.	Utiliza la obligación del objeto de referencia	Utiliza la máxima ocurrencia del objeto de referencia	Clase	92-95
91	Name	Nombre de la medida básica de calidad de los datos aplicada a los datos.	0	1	Cadena de caracteres	Texto libre
92	definition	Definición de la medida básica de calidad de los datos.	0	1	Cadena de caracteres	Texto libre
93	example	Ilustración de la utilización de una medida de la calidad de los datos.	Op	1	Clase	MeasureDescription << Datatype >> (C.2.2.5)
94	valueType	Tipo de valor para el resultado de la medida básica (debe ser uno de los tipos de datos definidos en la Norma ISO 19103:2015)	0	1	Clase	TypeName << interface >> (véase la Norma ISO 19103:2015)

C.2.2.4 Parámetro de calidad de los datos

En la tabla C.9 se describe el modelo UML que se muestra en la figura 10.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla C.9 – Parámetro de calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
95	MeasureParameter	Parámetro de calidad de los datos.	Utiliza la obligación del objeto de referencia	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Líneas 97-101
96	name	Nombre del parámetro de calidad de los datos.	0	1	Cadena de caracteres	Texto libre
97	definition	Definición del parámetro de calidad de los datos.	0	1	Cadena de caracteres	Texto libre
98	description	Descripción del parámetro de calidad de los datos.	Op	1	Clase	MeasureDescription << Datatype >> (C.2.2.5)
99	valueType	Tipo de valor para el parámetro de calidad de los datos (debe ser uno de los tipos de datos definidos en la Norma ISO 19103:2015).	0	1	Clase	TypeName << interface >> (véase la Norma ISO 19103:2015)
100	valueStructure	Estructura para el parámetro de calidad de los datos.	Op	1	Clase	ValueStructure << CodeList >> (C.3.3)

C.2.2.5 Descripción de la medida de calidad de los datos

En la tabla C.10 se describe el modelo UML que se muestra en la figura 10.

Tabla C.10 – Descriptor de la medida de la calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
101	MeasureDescription	Descripción de la medida de la calidad de los datos.	Utiliza la obligación del objeto de referencia	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Líneas 103-105
102	textDescription	Texto descriptivo.	0	1	Cadena de caracteres	Texto libre
103	extendedDescription	Ilustración.	Op	1	Clase	MD_Browse-Graphic (véase la Norma ISO 19115-1:2014, tabla B.17.3)
104	formula	Descripción de fórmulas para la medida de la calidad.	Op	N	Clase	FormulaType << Datatype >> (C.2.2.7)

C.2.2.6 Referencia de la fuente de la medida de la calidad de los datos

En la tabla C.11 se describe el modelo UML que se muestra en la figura 10.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla C.11 – Fuente de referencia de la medida de la calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
105	SourceReferen ce	Referencia a la fuente de la medida de la calidad de los datos.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Línea 107
106	citation	Referencia a la fuente.	0	1	Clase	CI_Citation << DataType >> (véase la Norma ISO 19115-1:2014, tabla B.16)

C.2.2.7 Tipo de fórmula de la calidad de los datos

En la tabla C.12 se describe el modelo UML que se muestra en la figura 10.

Tabla C.12 – Tipo de fórmula de la calidad de los datos

	Nombre/rol	Definición	Obligación/condición	Máxima ocurrencia	Tipo de datos	Dominio
107	FormulaType	Descripción de la fórmula.	Utiliza la obligación del objeto de referencia.	Utiliza la máxima ocurrencia del objeto de referencia.	Clase	Líneas 109-112
108	key	Explicación de la fórmula	0	1	Cadena de caracteres	
109	language	Lenguaje utilizado para expresar la fórmula.	0	1	Clase	FormulaLanguage << CodeList >> (C.3.3)
110	languageVersi on	Versión del lenguaje.	0	1	Cadena de caracteres	
111	mathematicalF ormula	Fórmula en el lenguaje elegido.	0	1	Cadena de caracteres	

C.3 Lista de códigos

C.3.1 Introducción

A continuación, se pueden encontrar las clases estereotipo << CodeList >>. Estas clases estereotipo no contienen las columnas “obligación/condición”, “máxima ocurrencia”, “tipo de datos” y “dominio”. Dado que una <<CodeList>> es extensible, ninguna de estas clases contiene el valor “otros”.

C.3.2 Tipo de método de evaluación

En la tabla C.13 se describe el modelo UML que se presenta en la figura 12.

Tabla C.13 – Tipo de método de evaluación

	Nombre	Definición
1	EvaluationMethodTypeCode	Tipo de método para evaluar una medida identificada de la calidad de los datos.
2	directInternal	Método para evaluar la calidad de un conjunto de datos basado en la inspección de ítems del conjunto de datos en el que todos los datos requeridos son internos al conjunto de datos evaluado.
3	directExternal	Método de evaluación de la calidad de un conjunto de datos basado en la inspección de ítems del conjunto de datos en el que se requieren datos de referencia externos al conjunto de datos evaluado.
4	indirect	Método de evaluación de la calidad de un conjunto de datos basado en conocimiento externo.

C.3.3 Estructura de valores

En la tabla C.14 se describe el modelo UML que se presenta en la figura 10.

Tabla C.14 – Estructura de valores

	Nombre	Definición
1	ValueStructure	Estructura del parámetro de la calidad de los datos
2	bag	Colección finita, sin orden, de ítems relacionados (objetos o valores) que puede contener duplicados (ISO 19103).
3	set	Colección sin orden de ítems relacionados (objetos o valores) sin repetición (ISO 19103)

4	sequence	Colección finita, ordenada de ítems relacionados (objetos o valores) que puede repetirse (ISO 19103)
5	table	Disposición de datos en la cual cada ítem puede identificarse por medio de argumentos o claves (ISO/IEC 2382).
6	matrix	Matriz rectangular de números (ISO/TS 19129)
7	coverage	Fenómeno que actúa como una función que devuelve valores de su rango para cualquier posición directa dentro de su dominio (ISO/FDIS 19123-1)

C.3.4 Lenguaje de fórmulas

En la tabla C.15 se describe el modelo UML que se presenta en la figura 10.

Tabla C.15 – Lenguaje de fórmulas

	Nombre	Definición
1	FormulaLanguage	Lenguaje de codificación de la fórmula
2	MathJSON	Formato basado en JSON para representar fórmulas matemáticas (w3c.github).
3	MathML	Lenguaje basado en XML para describir notación matemática (mozilla.org).
4	OpenMath	Estándar ampliable para representar la semántica de los objetos matemáticos (openmath.org)

**ANEXO D
(Informativo)
Evaluación e informe de la calidad de los datos**

D.1 Introducción

Este anexo ofrece un ejemplo principal que describe la evaluación y la presentación de informes sobre la calidad de los datos.

De manera adicional, se ofrecen otros ejemplos en el capítulo E.5, en este caso orientados a informar sobre metadatos de casos específicos tales como resultado descriptivo, metacalidad y evaluación por muestreo.

D.2 Descripción del conjunto de datos

D.2.1 Especificaciones de producto de datos

D.2.1.1 Generalidades

Las especificaciones de producto de datos definidas en los siguientes apartados describen el universo de discurso. Las especificaciones definen aquellos objetos geográficos, atributos y relaciones que se consideran importantes y que deberían estar presentes en el conjunto de datos.

NOTA: Este no es un ejemplo completo de una especificaciones de un producto de datos (véase Norma ISO 19131).

El producto estará compuesto por una red de transporte (camino y carreteras), construcciones (casas y edificios industriales) y árboles.

D.2.1.2 Tipos de objeto geográfico

Cada tipo de objeto geográfico, con cero o más atributos, se lista en la tabla D.1. Cada nombre de atributo está seguido por un tipo de valor (por ejemplo, lista de códigos, tipos de datos, cadena de caracteres o número entero) y, opcionalmente, por el dominio del valor del atributo.

D.2.1.2 Tipos de objeto geográfico

Cada tipo de objeto geográfico, con cero o más atributos, se lista en la tabla D.1. Cada nombre de atributo está seguido por un tipo de valor (por ejemplo, lista de códigos, tipo de datos, cadena de caracteres o número entero) y, opcionalmente, por el dominio del valor del atributo.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla D.1 – Tipos de objeto gráfico

	Tipo de objeto geográfico	Nombre de atributo	Tipo de valor	Dominio
Edificios	Edificio industrial			
	Casa	nombre de la familia	Cadena de caracteres	
		número de residentes	Entero	
Red de transporte	Camino			
	Carretera	estado	Lista de códigos: Pavimento	asfaltado, sin asfaltar
	Árbol	altura	Lista de códigos: TreeHeight (altura de árbol)	A: de 1 m a 3 m B: de 3 m a 5 m C: de 5 m a 10 m D: más de 10 m

D.2.1.3 Reglas

Los tipos de objeto geográfico de la tabla D.1 deben seguir las siguientes reglas:

- Los árboles con un altura menor de 1 m no deben registrarse;
- El atributo “estado” de una carretera puede no tener valor (“valor indeterminado”); aprobar;
- Los atributos “nombre” y “número de residentes” de una casa pueden no tener valor (“valor indeterminado”).

D.2.1.4 Niveles de conformidad de la calidad

Requisito general de la calidad de los datos: un conjunto de datos se debe considerar conforme solo si cumple los requisitos que se listan a continuación:

- a) Solamente los tipos de objeto geográfico y de atributo definidos en estas especificaciones de producto de datos pueden estar presentes en el conjunto de datos.

Red de transporte.

- b) Para cada tipo de objeto geográfico puede omitirse un máximo de 2 ítems.
- c) Para cada tipo de objeto geográfico puede existir 2 ítems en exceso como máximo.
- d) Cómo máximo dos instancias de objeto geográfico pueden estar clasificadas erróneamente como otro tipo de objeto geográficos de la red de transporte, y ninguna como otros tipos.

Edificios.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- e) Para cada tipo de objeto geográfico puede omitirse un máximo de 2 ítems.
- f) Para cada tipo de objeto geográfico puede existir 2 ítems en exceso como máximo.
- g) Como máximo dos instancias de objeto geográfico pueden clasificarse erróneamente como otro tipo de objeto geográfico de edificios, y ninguna como otros tipos.

Árboles.

- h) Máximo 10% de árboles omitidos.
- i) Máximo 10% de árboles en exceso.
- j) Máximo 20% de árboles con altura incorrecta.
- k) Ninguna instancia de objeto geográfico puede clasificarse erróneamente como otros tipos de objeto geográfico.

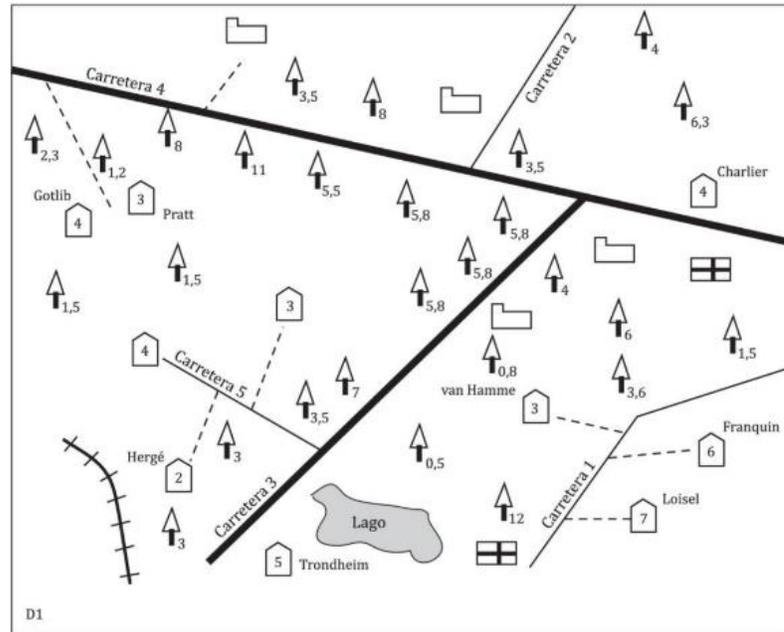
D.2.2 Representación del mundo real, del universo de discurso y del conjunto de datos

La relación entre las tres figuras de este apartado es la siguiente:

- a) La figura D.1 representa el “mundo real”, que generalmente contiene más objetos geográficos que los que contendrá el conjunto de datos;
- b) La figura D.2 representa el “universo de discurso” dado por las especificaciones de producto de datos. Es esa parte del mundo real la que tiene que estar incluida en el conjunto de datos, si el conjunto de datos se ha producido de forma completa y exacta;
- c) La figura D.3 representa el conjunto de datos tal y como se ha producido.

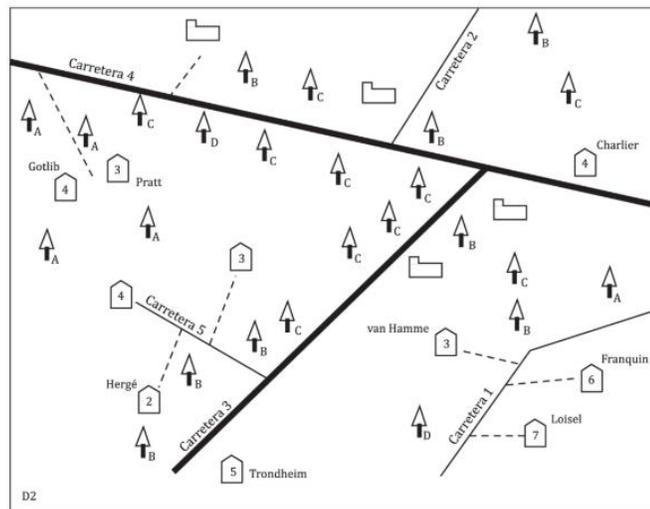
En todas las figuras:

- El dígito, o la letra que representa el dominio de los dígitos, debajo del símbolo del árbol es la altura del árbol en metros,
- El dígito dentro del símbolo de una casa es el número de habitantes de la casa,
- El nombre de los ocupantes de una casa se indica junto al símbolo de la casa.



- Leyenda
- | | | | |
|--|--------------------------|--|-------------------------|
| | Hospital | | Vía férrea |
| | Edificio industrial | | Casa con x ocupantes |
| | Árbol de 3,5 m de altura | | Carretera: asfaltada |
| | Camino | | Carretera: sin asfaltar |

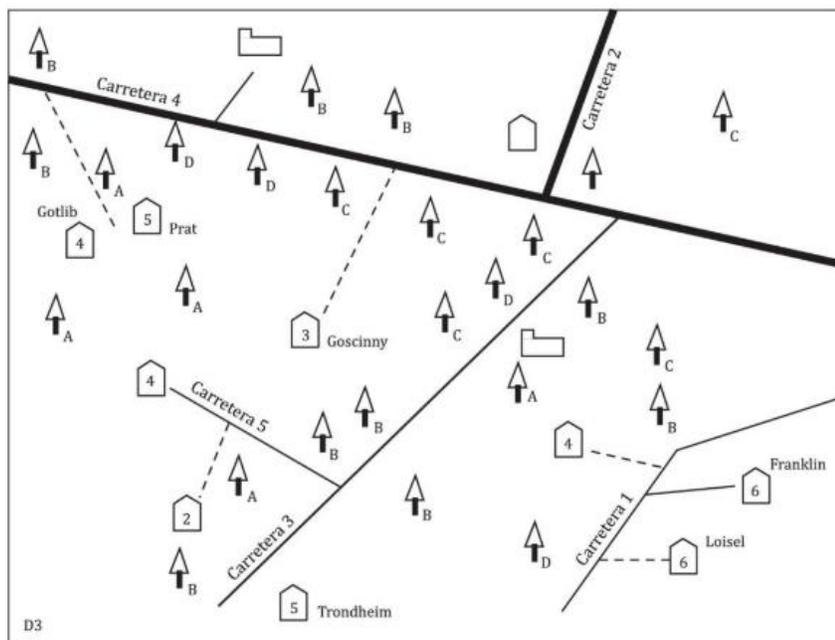
Figura D.1 – Representación gráfica del “mundo real”



- Leyenda
- | | | | |
|--|--------------------------------------|--|-------------------------|
| | Edificio industrial | | Casa con x ocupantes |
| | Árbol con altura de entre 5 m y 10 m | | Carretera: asfaltada |
| | Camino | | Carretera: sin asfaltar |

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Figura D.2 – Representación gráfica del universo de discurso



Leyenda

	Edificio industrial		Casa con x ocupantes
	Árbol con altura de entre 5 m y 10 m		Carretera: asfaltada
	Camino		Carretera: sin asfaltar

Figura D.3 – Representación gráfica del conjunto de datos

D.3 Proceso de evaluación de la calidad

D.3.1 Especificar las unidades de la calidad de los datos

- Una unidad de la calidad de los datos está compuesta por un ámbito y elemento(s) de la calidad. En este ejemplo se evalúa la conformidad de la completitud y la calidad temática frente a las especificaciones de producto de datos.
- La primera unidad de la calidad está compuesta por la coherencia conceptual, la completitud (comisión y omisión) y la corrección de la clasificación temática evaluada en las redes de transporte y los edificios.
- Otra unidad de la calidad está compuesta por la exactitud de los atributos cuantitativos evaluada en un tipo de objeto geográfico (árbol).

En el anexo F se ofrecen directrices para elegir los elementos de calidad de los datos apropiados.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

D.3.2 Especificar las medidas de calidad de los datos

Para describir la coherencia lógica se usa la siguiente medida:

Medida 9, “cumplimiento del esquema conceptual”

Para describir la completitud se usan las siguientes medidas:

Medida 1, “ítem en exceso”;

Medida 2, “número de ítems en exceso”;

Medida 3, “Índice de ítems en exceso”;

Medida 5, “ítem omitido”;

Medida 6, “número de ítems omitidos”;

Medida 7, “índice de ítems omitidos”;

Para describir la calidad temática se usa la siguiente medida:

Medida 62, “matriz de error de la clasificación”.

Para describir la conformidad general con las especificaciones se usa la siguiente medida:

Medida 101, “cumplimiento de especificaciones de producto”.

D.3.3 Especificar los procedimientos de evaluación de la calidad de los datos

Para este ejemplo se usa un procedimiento directo externo.

En este ejemplo se usa una inspección completa.

NOTA: En el apartado D.5.5 se describe un ejemplo de procedimiento de muestreo.

D.3.4 Especificar los niveles de conformidad de la calidad

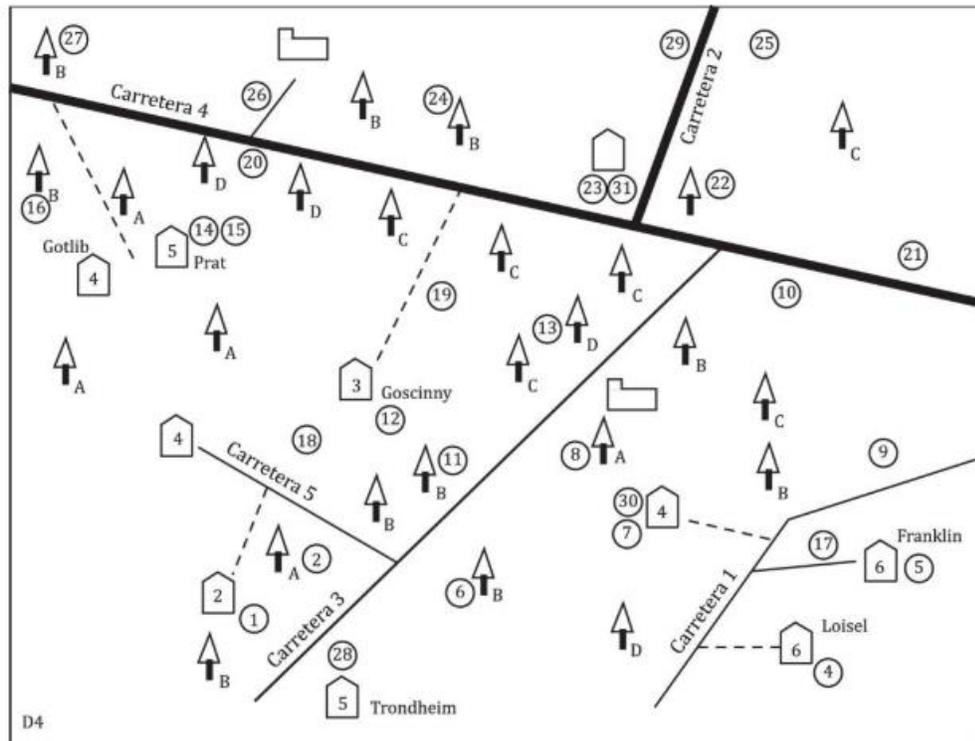
Se espera que el conjunto de datos sea conforme con los niveles de conformidad de la calidad especificados en el apartado D.2.1.4.

D.3.5 Determinar la salida de la evaluación de la calidad de los datos (resultado)

D.3.5.1 Identificación de errores

Mediante la comparación del conjunto de datos representado por la figura D.3 con el universo de discurso representado por la figura D.2, se puede generar una lista con los errores del conjunto de datos de ejemplo representado por la figura D.4.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.



Leyenda

	Edificio industrial		Número de error
	Árbol con altura de 3,5 m		Casa con x ocupantes
	Camino		Carretera: asfaltada
	Carretera: sin asfaltar		

Figura D.4 – Representación gráfica de la localización de los errores del conjunto de datos

La lista siguiente recoge los errores detectados y una numeración de esos errores como referencia:

Errores de omisión y comisión al registrar los árboles. Hay tres árboles (nº 6, nº 8, nº 27) en exceso y dos árboles (nº 9, nº 25) que faltan.

Errores de omisión y comisión al registrar los caminos. Hay un camino (nº 18) que falta y uno (nº 19) que está en exceso.

Hay una casa en el lugar de un edificio industrial (nº 23).

Dos caminos están mal codificados como carreteras (nº 17, nº 28).

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Falta una casa (n° 21).

Errores de atributo en las carreteras. En dos carreteras el atributo “estado” es erróneo (n° 19, n° 18).

En el conjunto de datos se han representado dos árboles de altura inferior a 1 m (n° 6, n° 8) como ítems en exceso.

Falta el código de clase en el atributo altura del árbol. A un árbol le falta el código de clase mientras que en el universo de discurso aparece clasificado como B (n° 22).

El atributo altura del árbol está mal clasificado. Seis árboles tienen clase de altura errónea (n° 2, n° 11, n° 13, n° 16, n° 20, n° 24).

Errores en el atributo de casa “nombre de la familia”. Las casas denominadas “van Hamme” (n° 7) y “Hergé” (n° 1) en el universo de discurso no tienen nombre asignado en el conjunto de datos. La casa denominada como “Goscinnny” en el conjunto de datos (n° 12) no tiene nombre en el universo de discurso.

Errores en el atributo de casa “nombre de la familia”. Las casas denominadas “Franquin” (n° 5) y “Pratt” (n° 15) en el universo de discurso se registran como “Franklin” y “Prat” en el conjunto de datos, respectivamente.

Errores en el atributo de casa del número de ocupantes. Falta este atributo en una casa (n° 31), y en tres casas está equivocado (n° 4, n° 14, n° 30),

Error de omisión en edificios industriales. Falta un edificio industrial (n° 10).

NOTA: La clasificación de los errores como completitud (omisión/comisión) o calidad temática es subjetiva. Por ejemplo, un error de clasificación de una casa como un edificio industrial podría considerarse alternativamente como un error de omisión de la casa y un error de comisión del edificio industrial.

D.3.5.2 Coherencia lógica

Solo están presentes en el conjunto de datos los tipos de objeto geográfico y atributos definidos en las especificaciones de producto de datos. Véase el resultado de conformidad para la coherencia conceptual en la tabla D.2.

Tabla D.2 – Resultado de conformidad para la coherencia conceptual

Ámbito	Elemento de calidad	Requisitos de conformidad	Número de evaluaciones	Recuentos Sí/No	Apto
Conjunto de datos	Coherencia conceptual	1) Solo pueden estar presentes en el conjunto de datos definidos en el	1 (no se detectan errores)	1/0	Sí

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

		esquema de aplicación		
--	--	-----------------------	--	--

D.3.5.3 Completitud

D.3.5.3.1 Generalidades

En este ejemplo la completitud se clasifica por la clase de objeto geográfico. Los tipos de medidas evaluadas son para la comisión y la omisión. Las tablas D.3 a D.5 muestran los resultados.

D.3.4.3.2 Resultado cuantitativo

La tabla D.3 presenta un modo de clasificar la completitud usando valores cuantitativos.

Tabla D.3 – Completitud según la clase de objeto geográfico

Clase de objeto geográfico	Número de instancias en el universo de discurso	Recuento de comisiones	Porcentaje de comisión ^a	Recuento de omisiones	Porcentaje de omisión ^b
Camino	7	1	14	1	43
Carretera ^c	5	2	40	0	0
Árbol	25	1	12	2	8
Edificio industrial	4	0	0	1	50
Casa	10	0	10	1	10

a Porcentaje de comisión = número de ítems incluidos/número de ítems en el universo de discurso × 100
 b Porcentaje de omisión = número de ítems omitidos/número de ítems en el universo de discurso × 100
 c Si existe poca diferencia entre una carretera sin pavimentar y un camino amplio, el error de 2 comisiones de carreteras podría ser un error de clasificación; véase el apartado D.3.5.4.2.

D.3.5.3.3 Resultado de conformidad derivado

La tabla D.4 presenta los resultados de conformidad derivados de los resultados cuantitativos.

Tabla D.4 – Conformidad de la completitud

Identificador de la evaluación	Elemento de la calidad	Medida e identificador de la medida	Tipo de objeto geográfico	Requisitos de conformidad	LCA ^a	Recuento de errores	Población	Apto
1	Comisión	Ítem en exceso (2)	Camino	c	2	1	7	Sí
2	Omisión	Ítem omitido (5)	Camino	b	2	3	7	No
3	Comisión	Ítem en exceso (2)	Carretera	c	2	2	5	Sí
4	Omisión	Ítem omitido (5)	Carretera	b	2	0	5	Sí
5	Comisión	Ítem en exceso (1)	Árbol	i	10 %	3	25	No
6	Omisión	Ítem omitido (5)	Árbol	h	10 %	2	25	Sí
7	Comisión	Ítem en exceso (2)	Edificio Industrial	f	2	0	4	Sí
8	Omisión	Ítem omitido (5)	Edificio industrial	e	2	2	4	Sí
9	Comisión	Ítem en exceso (2)	Casa	f	2	1	10	Sí
10	Omisión	Ítem omitido (5)	Casa	e	2	1	10	Sí

a LCA = límite de calidad de aceptación.

D.3.5.3.4 Resultado de conformidad agregado

Los resultados de conformidad respecto a las redes de transporte (caminos y carreteras) y edificios (industriales y casas) se muestran agregados en al tabla D.5 usando la siguiente regla: si uno de los resultados originales es “No” el resultado agregado será “No”. (Apto/no apto al 100%; véase el anexo G).

Tabla D.5 – Conformidad agregada de la completitud

Ámbito	Elemento de la calidad	Requisitos de conformidad	Número de evaluaciones e identificadores (véase la tabla D.4)	Recuentos sí/no	Apto
Red de transporte	Omisión	b) Máximo dos omisiones para cada tipo de objeto geográfico	2 (identificadores de evaluación b y d)	1/1	No
Red de transporte	Comisión	c) Máximo dos comisiones para cada tipo de objeto geográfico	2 (identificadores de evaluación a y c)	2/0	Sí
Edificios	Omisión	e) Máximo dos omisiones para cada tipo de objeto geográfico	2 (identificadores de evaluación h y j)	2/0	Sí
Edificios	Comisión	f) Máximo dos comisiones para cada tipo de objeto geográfico	2 (identificadores de evaluación g e i)	2/0	Sí

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

D.3.5.4 Calidad temática – corrección de la clasificación

D.3.5.4.1 Generalidades

La información sobre la completitud puede aclararse aún más con información sobre la calidad temática. Por ejemplo, dos de los tres caminos omitidos están de hecho clasificados como carreteras (véase la tabla D.6). Los resultados se muestran en las tablas D.6 a D.8.

D.3.5.4.2 Resultado cuantitativo

Una forma de presentar errores asociados a la calidad temática es mediante la medida “matriz de errores de clasificación”.

La tabla D.6 es una matriz de errores de clasificación que muestra los errores por clase de objeto geográfico. La matriz explica lo bien que se han clasificado las instancias del conjunto de datos. Los diferentes porcentajes deberían referirse siempre a la población en el conjunto de datos.

NOTA: Una matriz de errores de clasificación es una matriz cuadrada donde el elemento i,j corresponde a la cantidad clasificada como perteneciente a la clase j cuando realmente pertenece a la clase i .

Tabla D.6 – Matriz de errores de clasificación de los objetos geográficos

	Camino	Carretera	Árbol	Universo de discurso	Conjunto de datos	Total
Camino	4	2	0	0	0	6
Carretera	0	5	0	0	0	5
Árbol	0	0	23	0	0	23
Edificio industrial	0	0	0	2	1	3
Casa	0	0	0	0	9	9
Total	4	7	23	2	10	46

La discrepancia entre el total y el número de ítems en el universo de discurso y el conjunto de datos proviene de los ítems en exceso y omitidos.

D.3.5.4.3 Resultado derivado de conformidad

En la tabla D.7 se presentan los resultados de conformidad derivados de los resultados cuantitativos.

Tabla D.7 – Conformidad de la calidad temática

Identificador de la evaluación	Elemento de la calidad	Medida	Tipo de objeto geográfico	Requisitos de conformidad	LCA	Recuento de errores de clasificación	Apto
11	Corrección de la clasificación temática	Número de objetos geográficos clasificados incorrectamente	Camino	d	2	2	Sí
12	Corrección de la clasificación temática	Número de objetos geográficos clasificados incorrectamente	Carretera	d	2	0	Sí
13	Corrección de la clasificación temática	Número de objetos geográficos clasificados incorrectamente	Edificio Industrial	g	2	1	Sí
14	Corrección de la clasificación temática	Número de objetos geográficos clasificados incorrectamente	Casa	g	2	0	Sí
15	Corrección de la clasificación temática	Número de objetos geográficos clasificados incorrectamente	Árbol	k	0	0	Sí

D.3.5.4.4 Resultado agregado de conformidad

Los resultados de conformidad respecto a las redes de transporte (caminos y carreteras) y edificios (industriales y casas) se muestran agregados en la tabla D.8 usando la siguiente regla: si uno de los resultados originales es “No” el resultado agregado será “No”. (Apto/no apto al 100%; véase el anexo G).

Tabla D.8 – Conformidad de corrección de la clasificación agregada

Ámbito	Elemento de la calidad	Requisitos de conformidad	Número de evaluaciones e identificador (véase tabla D.7)	recuentos sí/no	Apto
Red de transporte	Corrección de la clasificación temática	d) Para cada tipo de objeto geográfico un máximo de dos instancias clasificadas erróneamente como otro tipo de red de transporte.	2 (evaluación nº 11 y 12)	2/0	Sí
Edificios	Corrección de la clasificación temática	7) Máximo dos instancias de objeto geográfico clasificadas erróneamente como otro tipo de edificio.	2 (evaluación nº 13 y 14)	2/0	Sí

D.3.5.5 Calidad temática – exactitud de atributos cuantitativos

D.3.5.5.1 Generalidades

El tipo de medida evaluada en este ejemplo es para la exactitud de atributo cuantitativo.

En la tabla D.9 solo se tiene en cuenta los objetos que tienen homólogo en el mismo tipo (“clase”) de objeto geográfico. Los resultados se muestran en las tablas D.9 y D.10.

D.3.5.5.2 Resultado cuantitativo

El atributo altura de árbol se muestra en la tabla D.9

Tabla D.9 – Matriz de errores de clasificación del atributo altura – altura de árbol

Universo de discurso	Conjunto de datos				Total
	clase A de 1 m a 3 m	clase B de 3 m a 5 m	clase C de 5 m a 10 m	clase D > de 10 m	
Clase A	3	1	0	0	4
Clase B	1	5	0	0	6
Clase C	0	2	6	2	10
Clase D	0	0	0	2	2
Total	4	8	6	4	22

A un árbol le falta el código de clase y por tanto no se contabiliza en la matriz de errores de clasificación. De este error podría informarse como error de coherencia de dominio.

D.3.5.5.3 Resultado derivado de la conformidad

La tabla D.10 presenta el resultado de conformidad derivado de los resultados cuantitativos.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla D.10 – Conformidad de la calidad temática

Elemento de la calidad	Medida e identificador de la medida	Tipo de objeto geográfico / atributo	Requisitos de conformidad	LCA	Recuento de errores	Población	Apto
Exactitud de atributo cuantitativo	Matriz de errores de clasificación (62)	Árbol/clase de altura	j	20 %	6	22	No

D.3.5.6 Conformidad con las especificaciones de producto de datos agregada

En la tabla D.11 se agregan todos los resultados de conformidad para edificaciones, red de transporte y árboles junto a los de conformidad con el esquema conceptual para proporcionar la conformidad a con las especificaciones de producto de datos siguiendo la medida registrada “cumplimiento de especificaciones de producto”.

Tabla D.11 – Conformidad con las especificaciones de producto de datos

Ámbito	Requisitos de conformidad	Número de evaluaciones	Recuentos si/no	Apto
Conjunto de datos	Para ser conforme con los requisitos de calidad de los datos, un conjunto de datos debe pasar todos los requisitos de calidad de los datos del esquema de aplicación.	11 requisitos	8/3 (no superados los requisitos 2, 9, 10)	Conjunto de datos NO conforme

D.4 Informe de la calidad de los datos

D.4.1 Presentación en forma de metadatos

D.4.1.1 Generalidades

En los apartados D.4.1.2 a D.4.1.4 se presentan ejemplos de cómo informar sobre los resultados de la calidad en forma de metadatos, tal y como se describe en este documento (capítulo 11 y anexo C) y en la Norma ISO 19115-1. De hecho, una instancia de MD_Metadata agrega una o más instancias de DataQuality.

En los ejemplos, a algunas instancias de clases (DataQuality y (QualityElement) se les ha dado un identificador (ID) de acuerdo con los principios de XML. Estos identificadores se utilizan para referir esas instancias dentro de otras clases.

D.4.1.2 Informe de comisiones

En la tabla D.12 se presenta un ejemplo de cómo informar sobre los resultados cuantitativos, del resultado derivado de conformidad y del resultado agregado de conformidad, para los tipos de objeto geográfico de red de transporte.

Para los otros tipos de objeto geográfico del conjunto de datos, el mecanismo para informar sobre estos resultados es similar.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla D.12 – Informe de comisiones como metadatos

Elemento XML	Ejemplo	Comentario
DataQuality		
scope: MD_Scope		
level: MD_ScopeCode	Conjunto de datos	Ámbito de esta unidad de la calidad de los datos.
QualityEvaluationReport: QualityEvaluationReportInformation		
reportReference: CI_Citation		Referencia y resumen del informe de evaluación de la calidad adjunto.
title: CharacterString	Presentación de informe de evaluación de la calidad, véase el apartado D.4.2.	
date: CI_Date		
date: Date	2010-07-05	
dateType: CI_DateTypeCode	Creation	
abstract: CharacterString	El informe de evaluación de la calidad adjunto a esta evaluación de la calidad proporciona más detalles sobre el método de derivación y agregación.	
report: Commission <i>id = quantitative_commission</i>		En esta instancia de comisión, el resultado cuantitativo de cada tipo de objeto geográfico se proporciona para la medida 2 (número de ítems en exceso).
measure: MeasureReference		
nameOfMeasure: CharacterString	Número de ítems en exceso	
measureIdentification: MD_Identifier		
code: CharacterString	2	
measureDescription: CharacterString	Número de ítems dentro del conjunto de datos que no deberían haber estado en el conjunto de datos.	
evaluation: FullInspection		
evaluationMethodType: EvaluationMethodTypeCode	directExternal	
evaluationMethodDescription: CharacterString	Comparar el recuento de ítems en el conjunto de datos frente al recuento de ítems en el universo de discurso.	

Elemento XML	Ejemplo	Comentario
result: QuantitativeResult		Para facilitar la lectura, aquí solo se informa de las comisiones de caminos y carreteras, pero debe informarse sobre cada tipo de objeto geográfico dado que el ámbito de la calidad es el conjunto de datos.
resultScope: MD_Scope		
level: MD_ScopeCode	featureType	
levelDescription: MD_ScopeDescription		
features: GF_FeatureType	Camino	
value: Record	0	
valueUnit: UnitOfMeasure	Ninguno	
result: QuantitativeResult		
resultScope: MD_Scope		
level: MD_ScopeCode	featureType	
levelDescription: MD_ScopeDescription		
features: GF_FeatureType	Carretera	
value: Record	2	
valueUnit: UnitOfMeasure	Ninguno	
report: Commission <i>id = conformance_commission</i>		En esta instancia de comisión, el resultado derivado de conformidad se proporciona para cada tipo de objeto geográfico para la medida 1 (ítem en exceso).
measure: MeasureReference		
nameOfMeasure: CharacterString	Ítem en exceso	
measureIdentification: MD_Identifier		
code: CharacterString	1	
measureDescription: CharacterString	Indicación de que un ítem está incorrectamente presente en los datos.	
evaluation: AggregationDerivation		
evaluationMethodType: EvaluationMethodTypeCode	indirecta	
evaluationMethodDescription: CharacterString	Derivación a partir del resultado cuantitativo	
derivedElement: Element	<i>quantitative_commission</i>	Referencia a los resultados originales.
result: ConformanceResult		Resultado derivado de conformidad para la comisión de camino.
resultScope: MD_Scope		
level: MD_ScopeCode	featureType	

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Elemento XML	Ejemplo	Comentario	
levelDescription: MD_ScopeDescription		Para facilitar la lectura, aquí solo se informa de las comisiones de caminos y carreteras, pero debe informarse sobre cada tipo de objeto geográfico dado que el ámbito de la calidad es el conjunto de datos.	
features: GF_FeatureType	Camino		
specification: CI_Citation			
title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem b de la lista).		
date: CI_Date			
date: Date	2010-07-05		
dateType: CI_DateTypeCode	Creation		
pass: Boolean	Verdadero		
result: ConformanceResult			Resultado derivado de conformidad para la comisión de carreteras. Para facilitar la lectura, aquí solo se informa de las comisiones de caminos y carreteras, pero debe informarse sobre cada tipo de objeto geográfico dado que el ámbito de la calidad es el conjunto de datos.
resultScope: MD_Scope			
level: MD_ScopeCode	featureType		
levelDescription: MD_ScopeDescription			
features: GF_FeatureType	Camino		
specification: CI_Citation			
title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem b de la lista).		
date: CI_Date			
date: Date	2010-07-05		
dateType: CI_DateTypeCode	Creation		
pass: Boolean	Verdadero	Resultado agregado de conformidad para red de transporte.	
DQ_DataQuality <i>id = agg_commission1</i>			
scope: MD_Scope		Ahora el ámbito queda definido por los tipos de objeto geográfico de la red de transporte, lo que implica que la unidad de la calidad de los datos ha cambiado. Esta es la razón por la que se ha creado una nueva instancia de DataQuality.	
level: MD_ScopeCode	FeatureType		
levelDescription: MD_ScopeDescription			
features: GF_FeatureType	Red de transporte (carreteras y caminos)		
report: Commission			

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Elemento XML		Ejemplo	Comentario
	evaluation: AggregationDerivation		Método de agregación.
	evaluationMethodType: EvaluationMethodTypeCode	indirecta	
	evaluationMethodDescription: CharacterString	Agregación de apto/no apto al 100 % del resultado de conformidad de las comisiones para carreteras y caminos	
	evaluationProcedure: CI_Citation		
	title: CharacterString	Anexo G	
	date: CI_Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	derivedElement: Element	<i>conformance_commission</i>	Referencia a los resultados originales.
	result: ConformanceResult		
	specification: CI_Citation		
	title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem b de la lista)	
	date: CI_Date		
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	Pass: Boolean	Verdadero	

D.4.1.3 Informe de corrección de la clasificación

La tabla D.13 presenta un ejemplo de cómo informar sobre los resultados derivados de conformidad y sobre el resultado agregado de conformidad para los tipos de objeto geográfico de edificios.

Para los otros tipos de objeto geográfico del conjunto de datos el mecanismo para informar sobre estos resultados es similar.

Tabla D.13 – Informe de la corrección de la clasificación como metadatos

Elemento XML	Ejemplo	Comentario
DataQuality		
scope: MD_Scope		
level: MD_ScopeCode	Conjunto de datos	Ámbito de esta unidad de la calidad de los datos.
QualityEvaluationReport: QualityEvaluationReportInformation		
reportReference: CI_Citation		Referencia y resumen del informe de evaluación de la calidad.
title: CharacterString	Informando como informe de evaluación de la calidad; véase el apartado D.4.2.	
date: CI_Date		
date: Date	2010-07-05	
dateType: CI_DateTypeCode	Creation	
abstract: CharacterString	El informe de evaluación de la calidad que se adjunta a esta evaluación de la calidad proporciona todos los resultados cuantitativos que no se proporcionan en los metadatos y más detalles sobre el método de derivación y agregación.	
report: ThematicClassificationCorrectness <i>id = conformance_classification</i>		En esta instancia de corrección de la clasificación, se proporciona el resultado derivado de conformidad para cada tipo de objeto geográfico usando la medida 60 (número de objetos geográficos clasificados incorrectamente).
measure: MeasureReference		
nameOfMeasure: CharacterString	Número de objetos geográficos clasificados incorrectamente.	
measureIdentification: MD_Identifier		
code: CharacterString	60	
evaluation: AggregationDerivation		
evaluationMethodType: EvaluationMethodTypeCode	Indirecta	
evaluationMethodDescription: CharacterString	Derivación a partir de resultados cuantitativos reportados en el informe de evaluación de la calidad.	
QualityEvaluationReportDetails: CharacterString	Los resultados cuantitativos originales están descritos en el apartado D.3.5.5.2 del informe de evaluación de la calidad.	Referencia a los resultados originales.
result: ConformanceResult		Resultado derivado de conformidad para la clasificación de edificios industriales.
resultScope: MD_Scope		
level: MD_ScopeCode	featureType	Intencionadamente el resultado cuantitativo original no se proporciona en los metadatos.
levelDescription: MD_ScopeDescription		

Elemento XML		Ejemplo	Comentario
	features: GF_FeatureType	Edificio industrial	Este resultado se describe en el informe de evaluación de la calidad.
	specification: CI_Citation		
	title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem g de la lista)	El atributo QualityEvaluationReportDetails proporciona la referencia precisa al resultado original dentro del informe de evaluación de la calidad.
	date: CI_Date		Para facilitar la lectura, aquí solo se informa sobre los edificios industriales y las casas, pero debe informarse sobre cada tipo de objeto geográfico dado que el ámbito de la calidad es el conjunto de datos.
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	explanation: CharacterString	El resultado cuantitativo original se proporciona en D.3.5.5.2 del informe de evaluación de la calidad.	
	pass: Boolean	Verdadero	
	result: DQ_ConformanceResult		Resultado derivado de conformidad para la clasificación de edificios industriales.
	resultScope: MD_Scope		El resultado cuantitativo original no se proporciona intencionadamente en los metadatos. Este resultado es descrito en el informe de evaluación de la calidad. El atributo QualityEvaluationReportDetails proporciona la referencia precisa al resultado original dentro del informe de evaluación de la calidad.
	level: MD_ScopeCode	featureType	
	levelDescription: MD_ScopeDescription		
	features: GF_FeatureType	Casa	
	specification: CI_Citation		
	title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem g de la lista).	
	date: CI_Date		
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	Para mayor legibilidad, aquí solo se informa sobre la clasificación de edificios industriales y casas, pero debe informarse sobre cada tipo de objeto geográfico dado que el ámbito de la calidad es el conjunto de datos.
	explanation: CharacterString	El resultado cuantitativo original se proporciona en el informe de evaluación de la calidad.	
	pass: Boolean	Verdadero	
DataQuality <i>id = agg_classification2</i>			Resultado agregado de corrección de la clasificación para edificios.
	Scope: MD_Scope		Ahora el ámbito queda definido por los tipos de objeto geográfico de edificios, lo que implica que la unidad de la calidad de los datos ha cambiado. Es la razón por la que se ha creado una nueva instancia de DataQuality.
	level: MD_ScopeCode	FeatureType	
	levelDescription: MD_ScopeDescription		
	features: GF_FeatureType	Edificios (edificios industriales y casas)	
	report: ThematicClassificationCorrectness		

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Elemento XML		Ejemplo	Comentario
	evaluation: AggregationDerivation		Método de agregación.
	evaluationMethodType: EvaluationMethodTypeCode	Indirecta	
	evaluationMethodDescription: CharacterString	Agregación apto/no apto al 100 % del resultado de conformidad de la corrección de la clasificación para edificios industriales y casas	
	evaluationProcedure: CI_Citation		
	title: CharacterString	Anexo G	
	date: CI_Date		
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	derivedElement: Element	<i>conformance_classification</i>	Referencia a los resultados originales.
	result: ConformanceResult		
	specification: CI_Citation		
	title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem g de la lista).	
	date: CI_Date		
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	pass: Boolean	Verdadero	

D.4.1.4 Informe de conformidad con las especificaciones de producto de datos

La tabla D.14 presenta un ejemplo de cómo expresar la conformidad con las especificaciones de producto de datos mediante la agregación de resultados para los diferentes requisitos.

Tabla D.14 – Informe de conformidad en forma de metadatos

Elemento XML	Ejemplo	Comentario
DataQuality		
scope: MD_Scope		
level: MD_ScopeCode	Conjunto de datos	
QualityEvaluationReport: QualityEvaluationReportInformation		Referencia y resumen del informe de evaluación de la calidad adjunto.
reportReference: CI_Citation		
title: CharacterString	Informando como informe de evaluación de la calidad; véase el apartado E.4.2.	
date: CI_Date		

Elemento XML	Ejemplo	Comentario
date: Date	2010-07-05	
dateType: CI_DateTypeCode	Creation	
abstract: CharacterString	El informe de evaluación de la calidad que se adjunta a esta evaluación de la calidad proporciona información completamente detallada sobre la evaluación aplicada y los resultados obtenidos.	
report: QualityElement		Se utiliza una combinación de elementos de la calidad para informar sobre la conformidad del conjunto de datos con las especificaciones de producto de datos.
measure: MeasureReference		
nameOfMeasure: CharacterString	Cumplimiento de especificaciones de producto.	
measureIdentification: MD_Identifier		
code: CharacterString	101	
measureDescription: CharacterString	Indicación de que se han cumplido todos los requisitos contemplados en las especificaciones de producto.	
evaluation: DAggregationDerivation		
evaluationMethodType: EvaluationMethodTypeCode	Indirecto	
evaluationMethodDescription: CharacterString	Agregación de apto/no apto al 100 % de cada resultado de conformidad expresado en las especificaciones de producto de datos.	
evaluationProcedure: CI_Citation		
title: CharacterString	Anexo G	
date: CI_Date		
date: Date	2010-07-05	
dateType: CI_DateTypeCode	Creation	

Elemento XML	Ejemplo	Comentario
QualityEvaluationReportDetails: CharacterString	Los resultados originales se describen en los apartados D.3.5.2, D.3.5.3.4, D.3.5.4.4 y D.3.5.5.3 del informe de evaluación de la calidad.	Referencia a los resultados originales en el informe de evaluación de la calidad (resultado de conformidad de la coherencia conceptual, resultado de conformidad de la exactitud de atributos cuantitativos para las alturas de árbol, etc.).
derivedElement: Element	<i>agg_commission1</i>	Referencia al resultado agregado de conformidad de ítem en exceso para la red de transporte, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado agregado de conformidad de ítem en exceso para edificios, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado de conformidad de ítem en exceso para árboles, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado agregado de conformidad de la omisión para red de transporte, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado agregado de conformidad de la omisión para edificios, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado de conformidad de la omisión para árboles, descrito anteriormente en los metadatos.
derivedElement: Element	<i>(id)</i>	Referencia al resultado agregado de conformidad de la corrección de la clasificación para red de transporte, descrito anteriormente en los metadatos.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Elemento XML		Ejemplo	Comentario
	derivedElement: Element	<i>agg_classification2</i>	Referencia al resultado agregado de conformidad de la corrección de la clasificación para edificios, descrito anteriormente en los metadatos.
	derivedElement: Element	<i>(id)</i>	Referencia al resultado de conformidad de la corrección de la clasificación para árboles, descrito anteriormente en los metadatos.
	result: DQ_ConformanceResult		
	specification: CI_Citation		
	title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4)	
	date: CI_Date		
	date: Date	2010-07-05	
	dateType: CI_DateTypeCode	Creation	
	explanation: CharacterString	3 de los 11 requisitos no se cumplen: el conjunto de datos no es conforme.	
	pass: Boolean	Falso	

D.4.2 Informe mediante un informe de evaluación de la calidad

La estructura del informe de evaluación de la calidad es libre.

D.5 Ejemplos adicionales

D.5.1 Generalidades

Algunos conceptos no se han descrito en el ejemplo anterior. Los ejemplos adicionales de los apartados D.5.2 a D.5.4 muestran como informar sobre resultados descriptivos, metacalidad y procedimientos de evaluación por muestreo.

Algunos conceptos no se han descrito en los ejemplos del capítulo D.4.

D.5.2 Informe de resultados descriptivos como metadatos

Algunas veces puede ser imposible expresar la evaluación de un elemento de calidad de los datos de una forma cuantitativa. En este caso podría utilizarse el resultado descriptivo. La tabla D.15 es un ejemplo de reporte de resultados descriptivos como metadatos.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla D.15 – Informe de resultados descriptivos como metadatos

Elemento XML	Ejemplo	Comentario
DataQuality		
scope: MD_Scope		
level: MD_ScopeCode	Conjunto de datos	El conjunto de datos describe objetos arqueológicos.
report: RelativeInternalPositionalAccuracy		
evaluation: IndirectEvaluation		
evaluationMethodType: EvaluationMethodTypeCode	Indirecto	
evaluationMethodDescription: CharacterString	Compara la exactitud posicional absoluta de los objetos arqueológicos y la exactitud posicional absoluta de los ríos.	
deductiveSource:CharacterString	Exactitud posicional de los ríos cercanos al campamento arqueológico.	
result: DescriptiveResult		
statement: CharacterString	La exactitud posicional relativa entre los objetos arqueológicos y los ríos es mayor que la exactitud posicional absoluta de los objetos arqueológicos (5 m)	

D.5.4 Informe de metacalidad en forma de metadatos

Se evalúa la exactitud posicional absoluta del levantamiento topográfico de un sitio arqueológico: el resultado es 5 m de exactitud.

Se ofrece una elevación de la calidad de la evaluación utilizando el elemento de metacalidad confianza, para el cual se utiliza una medida denominada “factor de seguridad”.

La tabla D.16 describe como informar sobre metacalidad como metadatos.

Tabla D.16 – Informe de metacalidad en forma de metadatos

Elemento XML	Ejemplo	Comentario
DataQuality		
scope: MD_Scope		
level: MD_ScopeCode	Conjunto de datos	
report: AbsolutExternalPositionalAccuracy id = <i>positionalaccuracy1</i>		Informe de exactitud posicional absoluta.
measure: MeasureReference		Se añade un identificador al elemento de la calidad de los datos para poder referirlo en el siguiente elemento de metacalidad.
nameOfMeasure: CharacterString	Valor eficaz del error planimétrico.	
measureIdentification: MD_Identifier		
code: CharacterString	47	Aquí no se han cumplimentado todos los atributos opcionales.
measureDescription: CharacterString	Desviación típica donde el valor verdadero no es estimado a partir de las observaciones dado que es conocido <i>a priori</i> .	
evaluation: FullInspection		
evaluationMethodType: EvaluationMethodTypeCode	directExternal	
evaluationProcedure: CI_Citation		
title: CharacterString	Procedimiento del IGN de evaluación de la calidad de los datos.	
date: CI_Date		
date: Date	1995-02-09	
dateType: CI_DateTypeCode	Creation	
result: QuantitativeResult		
value: Record	5	
valueUnit: UnitOfMeasure	Metro	
report: Confidence		El informe de metacalidad (confianza) en relación con el informe de exactitud previo.
relatedElement: Element	<i>positionalaccuracy1</i>	
measure: MeasureReference		
nameOfMeasure: CharacterString	Factor de seguridad	
measureIdentification: MD_Identifier		
code: CharacterString	1	
authority: CI_Citation		
title: CharacterString	Medidas del IGN	
date: CI_Date		
date: Date	1995-01-01	
dateType: CI_DateTypeCode	Creación	

Elemento XML	Ejemplo	Comentario
measureDescription: CharacterString	La razón entre la clase de exactitud de los elementos de la evaluación y la que se tiene que obtener en el conjunto de datos.	
evaluation: FullInspection		
evaluationMethodType: EvaluationMethodTypeCode	directExternal	
evaluationMethodDescription	Cuanto mayor es el "factor de seguridad" más fiable es la evaluación. El "factor de seguridad" tiene que ser mayor que 2 para validar la evaluación.	
evaluationProcedure: CI_Citation		
title: CharacterString	Arrêté 2003 (legislación francesa)	
date: CI_Date		
date: Date	2003	
dateType: CI_DateTypeCode	Publication	
result: QuantitativeResult		
value: Record	2.4	
valueUnit: UnitOfMeasure		

D.5.4.2 Ejemplo de informe calidad para un modelo líder de datos de alturas que utiliza texto libre

Información general sobre la calidad

La calidad medida o estimada se aplica a la totalidad de la superficie escaneada. Por lo que resulta importante, para el área en cuestión, consultar los metadatos para obtener información sobre la técnica de estudio, el tiempo de escaneo y la densidad de puntos. Un buen conocimiento de las condiciones locales del terreno también es valioso para la interpretación de los datos.

Compleitud

La densidad de puntos sobre el terreno varía según el tipo de terreno, el tipo de vegetación, la estación del año en que se realizó el escaneo como láser y una serie de otros factores. Esta variación significa que en algunas áreas la densidad de puntos es insuficiente mientras que en otras la densidad de puntos es alta. ^[34]

D.5.4.3 Ejemplo de informe de la calidad para la exactitud posicional de hidrografía

Los requisitos geométricos sobre la exactitud posicional dependen del grado en que se distinga el objeto es un área geográfica delimitada. Los objetos de hormigón tienen requisitos más estrictos que los objetos con límites difusos.

Los objetos tienen una exactitud posicional muy alta en el agua.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

La orilla se mide al nivel normal del agua, excepto en lagos y ríos regulados donde se mide a la línea de máximo embalse. La posición de la orilla puede variar debido a diferentes niveles del agua. La actualización solo se lleva a cabo cuando está claro que ha habido un cambio importante y duradero.

La exactitud posicional de los arroyos es alta en superficies abiertas, pero es variable en zonas forestales donde es difícil ver a través de la cubierta forestal. Los grandes errores con relación a los arroyos se están corrigiendo gradualmente utilizando datos de láser/altura en zonas forestales. Otros objetos hidrográficos tienen una exactitud posicional muy elevada.^[35]

D.5.5 Como informar sobre el procedimiento de muestreo

Este ejemplo se basa en una base de datos topográfica producida por un estudio topográfico de un país europeo. Los niveles de conformidad de calidad han sido definidos en las especificaciones de producto de datos.

El tipo de objeto geográfico carretera es evaluado en este ejemplo utilizando una evaluación por muestreo.

El procedimiento de muestreo se aplica usando los principios de la Norma ISO 2859-1, como se describe en la D.17.

Tabla D.17 – Procedimiento para muestreos

Etapa del proceso	Ejemplo
Definir un método de muestreo	Muestreo polietápico. Seleccionar suficientes unidades de muestreo para cubrir la fracción muestral. El muestreo está basado en la ponderación de objetos geográficos.
Definir los ítems	Todos los objetos geográficos.
Dividir el ámbito de la calidad de los datos (población) en lotes.	Número de conjuntos de datos.
Dividir los lotes en unidades de muestreo.	Número N de cuadrados de 1 km × 1 km.
Definir la fracción de muestreo o el tamaño de la muestra.	El tamaño del muestreo depende del valor LCA para ese lote.
Seleccionar las unidades muestrales.	Seleccionar el número requerido de ítems de unidades de muestreo para alcanzar la fracción de muestreo o el tamaño de muestra.
Inspeccionar los ítems en las unidades muestrales.	Inspeccionar cada ítem en las unidades de muestreo.

Si el requisito de calidad para el objeto geográfico es 1 no conformidad por 100 unidades (LCA = 1), entonces todos los objetos geográficos capturados se comprueban a partir de la fuente de datos. La inspección por muestreo se ejecuta cuando el LCA = 4 o 15.

Los lotes usados en las pruebas deberían consistir en conjuntos de datos producidos, en la medida de lo posible, al mismo tiempo y con los mismos métodos. Del lote, se selecciona un número N de unidades de muestro cuadradas de 1 km x 1 km para que el número de objetos geográficos de la muestra sea suficiente para un LCA = 4.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

La tabla D.18 es un ejemplo de cómo reportar la información del procedimiento de muestreo como metadatos.

Elemento XML	Ejemplo
DataQuality	
scope: MD_Scope	
level: MD_ScopeCode	Feature Type.
levelDescription: MD_ScopeDescription	
features: GF_FeatureType	Carretera.
report: DQ_Commission	
measure: MeasureReference	
nameOfMeasure: CharacterString	Número de ítems en exceso.
measureIdentification: MD_Identifier	
code: CharacterString	2
measureDescription: CharacterString	Número de ítems, dentro del conjunto de datos, que no deberían aparecer en el conjunto de datos.
evaluation: SampleBasedInspection	

Tabla D.18 – Informe de los resultados de la evaluación por muestreo como metadatos

evaluationMethodType: EvaluationMethodTypeCode	directExternal
evaluationMethodDescription: CharacterString	Muestreo polietápico. Seleccionar suficientes unidades de muestreo para alcanzar la fracción muestral. El muestreo está basado en objetos geográficos ponderados.
evaluationProcedure: CI_Citation	
title: CharacterString	Anexo E
date: CI_Date	
date: Date	2010-07-05
dateType: CI_DateTypeCode	Publication.
referenceDoc: CI_Citation	
title: CharacterString	ISO 2859-1
date: CI_Date	
date: Date	1999-11-18
dateType: CI_DateTypeCode	Publication
lotDescription: CharacterString	Un lote es un grupo de bases de datos (por ejemplo, hoja de mapa a escala 1:10000) que son tomadas para la inspección. El tamaño del lote es el número de objetos geográficos en el lote. Todas las carreteras en el conjunto de datos (un lote para el conjunto de datos completo)
samplingScheme: CharacterString	Del lote, se muestrea un área de un cierto número de cuadrados de 1 km × 1 km, de forma que el número de carreteras en la muestra sea al menos el mismo que el requerido para un LCA = 4.
samplingRatio: CharacterString	En promedio se recomienda como tamaño práctico de lote un área que comprenda hojas formateadas (16 bases de datos) con 6 a 10 cuadrados (de 1 km × 1 km).
result: ConformanceResult	
specification: CI_Citation	
title: CharacterString	Especificaciones de producto de datos (véase D.2.1.4, ítem b de la lista).
date: CI_Date	
date: Date	2010-07-05
dateType: CI_DateTypeCode	Creation
Pass: Boolean	Verdadero

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

**ANEXO E
(informativo)**

Métodos de muestreo para la evaluación de la calidad de los datos

E.1 Introducción

Este anexo proporciona pautas para definir muestras e idear métodos de muestreo. En los muestreos para evaluar la conformidad frente a unas especificaciones de producto de datos se pueden aplicar la serie de Normas ISO 2859 y la Norma 3951. Originalmente, estos documentos fueron desarrollados para un uso no espacial. Este anexo describe cómo aplicar la serie de Normas ISO 2859 y la Norma ISO 3951-1, y otras técnicas de muestreo espacial a los datos geográficos.

E.2 Lote e ítem

Lote e ítem son dos conceptos importantes en los métodos de inspección por muestreo especificados en la serie de Normas ISO 2859 y en la Norma ISO 3951-1. Un lote es la mínima unidad para la que la calidad puede evaluarse. Un ítem es la mínima unidad para inspeccionar y el reproductor de datos debería definirlo de acuerdo con las especificaciones de producto de datos.

E.3 Tamaño de la muestra

El tamaño de una población, y consecuentemente el tamaño de las muestras, puede definirse de acuerdo con diferentes supuestos base sobre los ítems. La definición de un tamaño de muestra requiere una indicación explícita de los ítems. La tabla E.1 presenta ejemplos de diferentes supuestos base.

La diferencia entre las perspectivas se ilustra en la figura E.1. La figura en su conjunto representa los datos dentro del ámbito de la calidad de los datos. En la figura se marca una posible área de muestreo de aproximadamente el 15% del total del área del ámbito de la calidad de los datos, pero en este caso solo se encuentran dentro de esa área alrededor del 10% de la longitud de las curvas y un 0% de los vértices.

NOTA 1: El ámbito de la calidad de los datos es el área del marco exterior. El área de muestreo es el rectángulo sombreado.

Para salvar las dificultades de muestreo como las comentadas sobre la figura E.1, el tamaño y localización de la muestra puede definirse usando una combinación de diferentes criterios, reforzando de esta manera la representatividad de la muestra.

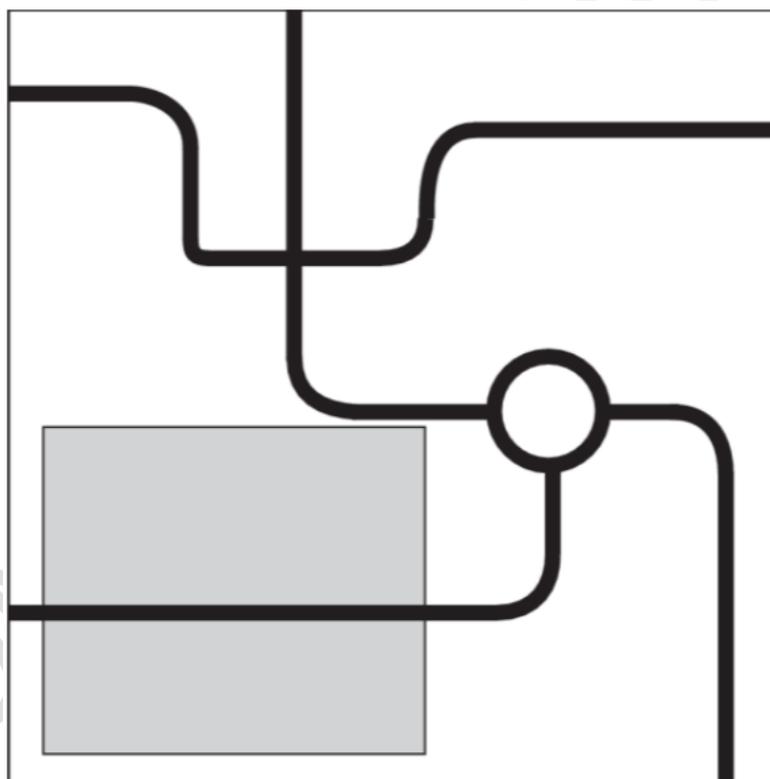
NOTA 2: El ámbito de la calidad de los datos es el área del marco exterior. El área de muestreo es la del rectángulo sombreado.

EJEMPLO: Se recomienda que la muestra incluya un 10% del área cubierta por el conjunto de datos y que contenga no menos de un 5% de la longitud total de las curvas que describen los objetos del conjunto de datos.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Tabla E.1 – Diferentes supuestos base para definir una población

Supuesto base	Tamaño del conjunto de datos	Tamaño de la muestra
Objetos geográficos	Número de objetos geográficos de un tipo determinado.	Número de objetos geográficos de un tipo determinado, expresado como un porcentaje del total de objetos.
Área cubierta	Área cubierta por el conjunto de datos.	Área cubierta por la muestra expresada como un porcentaje del área total.
Curvas	Longitud total de las curvas en el conjunto de datos.	Longitud de las curvas muestreadas expresadas como porcentaje de la longitud total.
Vértices	Número total de vértices que describen curvas o áreas en el conjunto de datos.	Número de vértices en la muestra expresado como un porcentaje del número total de vértices.



NOTA: El ámbito de la calidad de los datos es el área del marco exterior. El área de muestreo es la del rectángulo sombreado.

Figura E.1 – Efecto de la localización del área de muestreo sobre la representatividad de los ítems en la muestra

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

E.4 Estrategias de muestreo

E.4.1 Introducción

Teniendo en consideración los aspectos específicos de los datos geográficos, este apartado proporciona una guía para definir muestras y métodos de muestreo. Las estrategias de muestreo que se describen en este anexo se muestran gráficamente en la figura E.2. Hay dos aspectos a considerar en una estrategia de muestreo: los ítems a muestrear (área u objeto geográfico) y la manera de seleccionar los ítems (por probabilidad o por criterios).

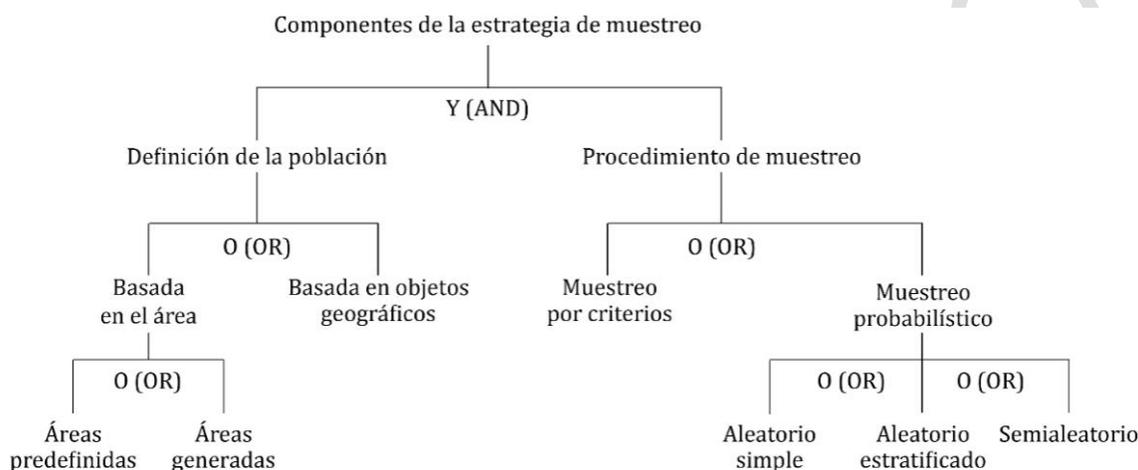


Figura E.2 – Relaciones entre estrategias de muestreo

E.4.2 Muestreo probabilístico frente a muestreo basado en criterios

E.4.2.1 Diferencias

El muestreo probabilístico aplica la teoría de muestreo y conlleva una selección aleatoria de los ítems de la muestra. La característica esencial de un muestreo probabilístico es que cada uno de los miembros de la población de la que se selecciona la muestra tiene una probabilidad conocida de ser seleccionado. Cuando se usa el muestreo probabilístico, se pueden realizar inferencias estadísticas sobre la población muestreada. Los diseños de muestras basadas en criterios conllevan la selección de las muestras a partir del conocimiento experto o juicio profesional.

E.4.2.2 Muestreo aleatorio simple

El muestreo aleatorio simple se basa en la probabilidad y supone la selección aleatoria de las muestras. La muestra (por ejemplo, objetos geográficos, localización, tiempo) se determina usando números aleatorios para identificar los ítems y todas las selecciones posibles son igualmente probables. El muestreo aleatorio simple es útil cuando la población de interés es relativamente homogénea respecto a las características a muestrear, es decir, sin patrones notables ni agrupamientos. Este método puede generar

Este documento comprende un borrador de Proyecto de Norma Técnica diseñado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

un resultado que no cubra de manera representativa el área, es decir, es posible que la muestra seleccionada proceda solo una parte del área.

E.4.2.3 Muestreo aleatorio estratificado

El muestreo estratificado requiere que la población sea separada en estratos o subpoblaciones que no se solapen, de forma que los ítems de una muestra son más homogéneos en un mismo estrato que los ítems de una muestra de diferentes estratos. Para una misma población, esta estrategia de muestreo tiene el potencial de alcanzar una mayor precisión en la estimación de la medida y varianza que una estrategia no estratificada.

E.4.2.4 Muestreo semialeatorio

El muestreo semialeatorio o sistemático aplica la selección aleatoria de los ítems de muestreo iniciales (por ejemplo, localización, tiempo u objeto geográfico) y reglas para la selección de todos los restantes ítems. Un ejemplo de muestreo semialeatorio o sistemático es una malla de muestreo donde la posición inicial de la malla se determina aleatoriamente y las muestras se toman a intervalos de espacio regulares (con celdas de la malla). El muestreo sistemático en malla se utiliza para la búsqueda de agrupaciones y para inferir medidas, percentiles y otros parámetros, y es útil para estimar tendencias espaciales o patrones. Este método proporciona una forma práctica y fácil de asegurar la cobertura de un área.

E.4.3 Muestreo basado en objetos geográficos frente al muestreo basado en áreas

E.4.3.1 Muestreo basado en objetos geográficos (muestreo no espacial)

En una estrategia de muestreo basada en objetos geográficos, se seleccionan los ítems de la muestra según los atributos no espaciales de los objetos y no por su situación espacial. Dentro de un ámbito de la calidad, la muestra puede elegirse aleatoriamente asumiendo características de producción homogéneas para todo el ámbito de la calidad. En algunos casos puede tal vez requerirse una distribución homogénea de las muestras y el muestreo aleatorio simple potencialmente producir una muestra no satisfactoria porque quizás solo existe homogeneidad en subconjuntos; por ejemplo, patrones globales o agrupamientos que tengan lugar en la características a muestrear. En estos casos un muestreo estratificado o semialeatorio puede potencialmente proporcionar mejores resultados.

NOTA: Si el métodos de muestreo se define mediante una selección aleatoria de objetos geográficos, existe el riesgo de que se seleccione una muestra que se concentre en un área reducida (lo cual puede potencialmente no ser aceptable).

El muestreo semialeatorio puede utilizarse para asegurar la verificación de diferentes criterios sobre el tamaño de muestra o localización para cumplir limitaciones suplementarias sobre las muestras o para reducir los costes del proceso de inspección.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

EJEMPLO: Una compañía energética necesita evaluar la corrección de los atributos capturados para diferentes tipos de objetos geográficos. Se consideraron dos métodos: una selección aleatoria y una selección semialeatoria (seleccionando aleatoriamente los objetos geográficos de un tipo y luego capturando los objetos geográficos de los diferentes tipos restantes en la vecindad de los premios hasta cubrir los tamaños de muestra de cada tipo) que conduzca a una reducción de costes de inspección.

E.4.3.2 Muestreo basado en áreas (muestreo espacial)

En una estrategia de muestreo basada en áreas, la selección de las unidades de muestreo se basa en consideraciones espaciales. Las unidades de muestreo pueden ser áreas geográficas existentes (por ejemplo, áreas políticas o estadísticas) o cualquier otra participación del universo de discurso para la que se realiza el muestreo. Este tipo de muestreo puede usarse como la primera etapa del muestreo, seguido por un muestreo basado en objetos geográficos dentro de cada una de la subáreas.

EJEMPLO: Una selección aleatoria de áreas de la cuadrícula UTM de 1 km x 1 km para evaluar los atributos de los objetos contenidos en esa área.

La figura E.3 ilustra el resultado de la definición de las áreas a someter a inspección, que se han obtenido mediante la generación aleatoria de las coordenadas del centro de unos rectángulos de superficies iguales (con la limitación de no superponerse).

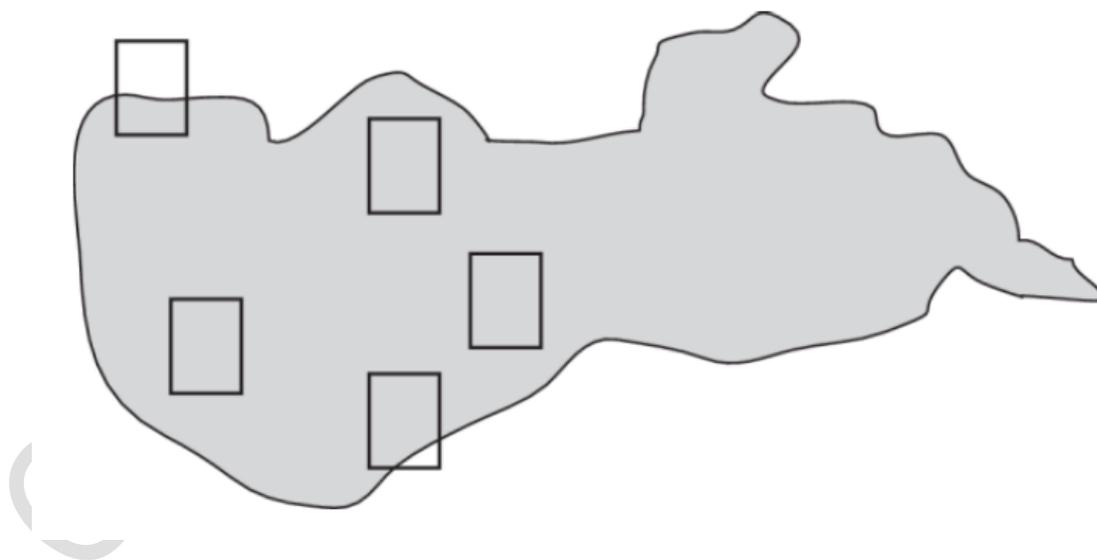
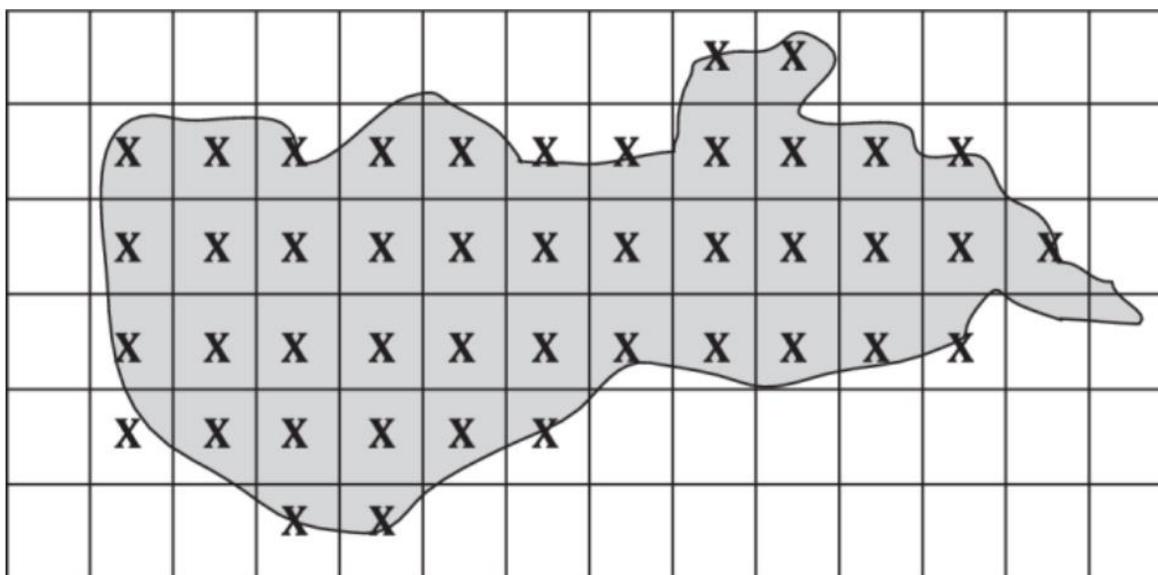


Figura E.3 – Ejemplo de muestreo aleatorio basado en áreas

Cuando es importante cubrir toda el área, entonces la situación de las muestras debería determinarse de acuerdo con un patrón regular o semirregular.

La figura E.4 ilustra un ejemplo de un muestreo semialeatorio (sistemático) con los objetos geográficos muestreados distribuidos según un patrón regular usado para evaluar la exactitud posicional del conjunto de datos.



NOTA: El símbolo “X” denota las celdas de la malla seleccionadas mediante una regla para su inclusión en la muestra.

Figura E.4 – Ejemplo de muestreo regular y no aleatorio basado en áreas

Si la distribución de los objetos geográficos no es homogénea, en el caso de muestreos semialeatorios se puede necesitar una partición espacial del conjunto de datos con tamaños distintos en áreas diferentes. Cuando se utiliza una allá de tamaño constante se necesita una regla para incluir o excluir celdas que no estén completamente incluidas dentro del área de interés.

E.5 Muestreos de base probabilística

E.5.1 Consideraciones generales

Cuando se aplican muestreos se deben tener en cuenta las siguientes consideraciones:

- a) Las áreas cubiertas por un conjunto de datos geográficos pueden potencialmente formar un espacio continuo. Cuando se divide el conjunto de datos en lotes se debería prestar especial atención a la omisión o comisión de ítems que cruzan los límites de los lotes.
- b) La calidad de los datos geográficos puede verse afectada por una variedad de factores, incluyendo la calidad de los datos origen y las habilidades de los operarios. El productor de datos debería tener el cuidado de definir muchos factores para alcanzar homogeneidad en términos de calidad.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Dependiendo del diseño del muestreo, pueden utilizarse ponderaciones de diseño (que en un muestreo probabilístico corresponde al inverso de la probabilidad de ser incluido en la muestra) para definir ítems en un lote de muestras.

E.5.2 Normas existentes para la inspección por muestreo

E.5.2.1 Generalidades

Basándose en las características de la producción y de acuerdo con las especificaciones de producto de datos, del conjunto de normas internacionales se deberían elegir los documentos adecuados para la inspección por muestreo. La Normas ISO 2859-1 se aplica principalmente a la inspección de series continuas de lotes. La Norma ISO 2859-1 se aplica principalmente a la inspección de series continuas de lotes. La Norma ISO 2859-2 puede aplicarse a lotes individuales aislados, mientras que la Norma ISO 2859-3 puede aplicarse a procedimientos de muestreo por lotes salteados. La Norma ISO 3951-1 puede aplicarse para la inspección por variables mediante porcentajes de ítems no conformes.

El nivel de conformidad de la calidad de un conjunto de datos se especifica como LCA (límite de calidad de aceptación) en la Norma ISO 3534-2.

NOTA: En ediciones previas de las Normas ISO 2859-1, ISO 2859-3 e ISO 3951-1, el LCA (límite de calidad de aceptación) en la Norma ISO 3534-2.

Para determinar la conformidad de cada ítem se deberían establecer límites de especificación cuando se aplica la serie de Normas ISO 2859, basada en las especificaciones de producto de datos. En la aplicación de la Norma ISO 3951-1, las estadísticas de calidad deberían establecerse basándose en las especificaciones de producto de datos.

E.5.2.2 Tablas útiles basadas en estas normas internacionales. Tamaño de la muestra y límites de rechazo

E.5.2.2.1 Visión general

Cuando se usan muestreos, el índice estimado de omisiones no puede compararse directamente con el LCA. La tabla E.2 y la tabla E.4 proporcionan directrices sobre el tamaño de la muestra de acuerdo con el tamaño del conjunto de datos y al nivel de rechazo asociado.

E.5.2.2.2 Evaluación de ítems conformes/no conformes con muestras

La tabla E.2 presenta el tamaño de la muestra recomendado de acuerdo con el tamaño de la población y el límite de rechazo asociado, para evaluar ítems conformes/no conformes, por ejemplo, para evaluar la completitud. Esto se basa en la distribución hipergeométrica. Se supone que las desviaciones se ajustan a esa distribución.

Cómo usar la tabla:

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- a) Decidir el tamaño de la población de los ítems que comprobar;
- b) Seleccionar el tamaño de la muestra, n , de la tabla;
- c) Llevar a cabo la evaluación y contar con el número de fallos,
- d) La población completa es rechazada si el número de fallos es igual o superior al límite de rechazo para los valores de n y p_0 (LCA).

Tabla E.2 – Valores estadísticos para comprobar el número de ítems conformes/no conformes

Nivel de significación 95%

Tamaño de población		$p_0 =$	0,5 %	1,0 %	2,0 %	3,0 %	4,0 %	5,0 %
Desde	A	Tamaño muestra (n)	Límite de rechazo					
1	8	Todos	1	1	1	1	1	1
9	50	8	1	1	1	2	2	2
51	90	13	1	1	2	2	2	3
91	150	20	1	2	2	3	3	4
151	280	32	1	2	3	3	4	4
281	400	50	2	3	3	4	5	6
401	500	60	2	3	4	5	6	7
501	1 200	80	3	3	5	6	7	8
1 201	3 200	125	3	4	6	8	10	11
3 201	10 000	200	4	6	8	11	14	16
10 001	35 000	315	5	7	12	16	20	23
35 001	150 000	500	6	10	16	23	28	34
150 001	500 000	800	9	14	24	33	42	51
> 500 000		1 250	12	20	34	49	63	76

Si el tamaño de muestra es el mayor que el tamaño mínimo mostrado en la tabla, el límite de rechazo debería calcularse individualmente. Esta prueba es válida para situaciones en las que la evaluación de la calidad está basada en una evaluación apto/no apto de los ítems.

NOTA: Existen otros rangos de valores estadísticos además de los presentados en la tabla E.2.

EJEMPLO: Comprobar las omisiones de casas (completitud/omisión) en un área definida.

Primero se selecciona un área de muestreo y se comprueba cada casa del área de muestreo para decidir si está presente en el conjunto de datos o no. Luego, se estiman (por conteo) el número de casas omitidas y el número total de casas. La pregunta que

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

se hace es: ¿el resultado es significativamente mayor que el LCA? Si es así, el conjunto de datos puede rechazarse. Si no es así, el conjunto de datos se acepta.

El conjunto de datos para comprobar consiste en 2440 edificios.

El tamaño de muestra (de la tabla E.2) es $n = 125$. El trabajo de comprobación en campo indica que 2 construcciones se han omitido, dando un índice de omisión estimado de: $2/(125 + 2) \times 100 = 1,6$ (donde la unidad es %).

LCA (de las especificaciones de producto de datos para el conjunto de datos, medido en %) es $p_0 = 0,5$.

1,6% es mayor que 0,5%; por lo tanto, se utilizará una prueba estadística para evaluar si los datos pueden rechazarse. Como se ha utilizado un muestreo el índice estimado de omisiones no puede compararse directamente con el LCA. Para ello se lleva a cabo un contraste de hipótesis unilateral y la tabla E.2 sirve de ayuda. El nivel de rechazo ($n = 125$, $p_0 = 0,5\%$) es 3. En la comprobación de campo se encontraron 2 ítems omitidos.

Conclusión: como 2 es menor que 3 (límite de rechazo) el conjunto de datos no puede rechazarse y se acepta.

E.5.2.2.3 Desviación típica

La tabla E.4 presenta el tamaño de muestra recomendado de acuerdo con el tamaño de población y el límite de rechazo asociado, cuando se mide una desviación típica.

Este método estadístico puede utilizarse para decidir si la desviación típica estimada para un tamaño de muestra es significativamente mayor que el LCA. La tabla E.4 se basa en la distribución normal y supone una distribución normal de las desviaciones.

Los símbolos y fórmulas de la tabla E.4 e muestran en la tabla E.3.

Tabla E.3 – Símbolos y fórmulas

Tamaño de población		$p_0 =$	0,5 %	1,0 %	2,0 %	3,0 %	4,0 %	5,0 %
Desde	A	Tamaño muestra (n)	Límite de rechazo					
1	8	Todos	1	1	1	1	1	1
9	50	8	1	1	1	2	2	2
51	90	13	1	1	2	2	2	3
91	150	20	1	2	2	3	3	4
151	280	32	1	2	3	3	4	4
281	400	50	2	3	3	4	5	6
401	500	60	2	3	4	5	6	7
501	1 200	80	3	3	5	6	7	8
1 201	3 200	125	3	4	6	8	10	11
3 201	10 000	200	4	6	8	11	14	16
10 001	35 000	315	5	7	12	16	20	23
35 001	150 000	500	6	10	16	23	28	34
150 001	500 000	800	9	14	24	33	42	51
> 500 000		1 250	12	20	34	49	63	76

El conjunto de datos no es suficientemente bueno (es decir, puede rechazarse con un 95% de significación) si la desviación típica estimada dividida entre el valor de F (tomado de la tabla E.4) es mayor que el LCA.

Tabla E.4 – Valores estadísticos para comprobar la desviación típica; nivel del significado del 95%

Tamaño de población		Tamaño de muestra (n)	$\sqrt{F_{0,05,n-1,\infty}}$
Desde	A		
26	50	5	1,54
51	90	7	1,45
91	150	10	1,37
151	280	15	1,30
281	400	20	1,26
401	500	25	1,23
501	1 200	35	1,20
1 201	3 200	50	1,16
3 201	10 000	75	1,13
10 001	35 000	100	1,12
35 001	150 000	150	1,09
150 001	500 000	200	1,08
> 500 000		200	1,08

EJEMPLO: Evaluación de la exactitud posicional/exactitud absoluta de tapas de registro.

De un conjunto de datos que contiene 450 tapas de registro, se ha determinado la posición de 25 tapas de registro (tamaño de muestra $n = 25$). La desviación típica estimada es $s = 21$ cm, y el LCA = 19 cm.

El límite inferior para el intervalo de confianza = $21 \text{ cm}/1,23$ (de la tabla E.4) = 1,71 cm.
El LCA = 19 cm.

Conclusión: La desviación típica del control no es significativamente mayor que la LCA y el conjunto de datos no puede rechazarse.

E.5.3 El proceso de muestro

E.5.3.1 Definición de ítems

Los ítems deberían definirse de acuerdo con las especificaciones de producto de datos o a sus requisitos. Si los ítems no conformes poseen una alta correlación estadística, son tratados como un solo ítem.

E.5.3.2 Definición de ámbitos de la calidad del conjunto de datos para inspeccionar

Si el ámbito de la calidad de los datos no es homogéneo, debería dividirse en subconjuntos homogéneos. Estos subconjuntos homogéneos deberían considerarse como ámbitos separados.

Se puede deducir homogeneidad cuando ocurren las siguientes condiciones:

- Los datos origen de la producción tienen casi la misma calidad.
- Los sistemas de producción (*hardware*, *software*, habilidad del operario) son esencialmente los mismos.
- Otros factores que pueden afectar potencialmente a la probabilidad de ocurrencia de no conformidades, como la complejidad y la densidad de objetos geográficos, son esencialmente los mismos.

E.5.3.3 División del ámbito de la calidad de los datos en lotes

Los lotes se generan dividiendo el ámbito de la calidad de los datos. Cuando exista una fuerte autorrelación espacial positiva en la ocurrencia de no conformidades es preferible reducir el tamaño de los lotes.

E.5.3.4 División del lote en unidades de muestreo

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Una unidad de muestreo puede ser un área geográfica existente o cualquier otra partición del universo de discurso a la que se dirige la inspección. Cuando la unidad de muestreo es un área geográfica se deberían proporcionar reglas para la consideración de ítems parcialmente incluidos en una unidad de muestreo.

E.5.3.5 Selección de unidades de muestreo para la inspección por muestreo aleatorio simple

El número total de ítems que pertenecen a las unidades de muestreo seleccionadas debería ser el especificado por los documentos pertinentes.

NOTA: Si los lotes son estadísticamente heterogéneos, el muestreo aleatorio simple no se puede aplicar con el mismo nivel de muestreo. Adicionalmente, la serie de normas ISO 2859 permite muestreos estratificados.

E.5.3.6 Inspección de las unidades de muestreo seleccionadas

Se inspeccionan todos los ítems que pertenecen a las unidades de muestreo seleccionadas. Se comparan con ítems en el conjunto de datos con el universo de discurso de acuerdo con la medida de la calidad elegida.

ANEXO F (informativo)

Métodos de muestreo para la evaluación de la calidad de los datos

F.1 Visión general

En algunos casos, puede haber varios elementos de calidad posibles para un requisito de calidad específico y un error detectado en una evaluación de la calidad. Este anexo ofrece directrices sobre qué elemento de calidad usar.

NOTA: Los elementos de calidad están descritos en el apartado 8.3.

F.2 Elemento de calidad de los datos

F.2.1 Generalidades

El apartado 8.3 define las siguientes categorías diferentes de elementos de calidad:

- Completitud (8.3.2);
- Coherencia lógica (8.3.3);
- Exactitud posicional (8.3.4);
- Calidad temática (8.3.5);
- Calidad temporal (8.3.6);

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

De los elementos restantes, la coherencia lógica es el único que puede evaluarse completamente sin conocimiento de la verdad terreno. Los requisitos de coherencia lógica y los procedimientos de evaluación manejan las “relaciones internas” entre los datos y cómo los datos se ajustan a las reglas establecidas en las especificaciones.

Las tres categorías completitud, exactitud posicional y calidad temática se usan para describir cómo se relaciona el conjunto de datos con el universo de discurso.

La calidad temporal consiste en una mezcla de elementos de calidad de los datos que depende en parte de reglas lógicas (comparable a la coherencia lógica) y en parte necesita conocimiento de la verdad terreno para ser evaluada (de manera similar a las categorías de completitud y las de exactitud).

F.2.2 Otros elementos de la calidad candidatos

Todos los elementos de calidad de los datos aceptados pueden evaluarse y el resultado de la evaluación no queda obsoleto simplemente porque el tiempo pase. Unos posibles candidatos “actualidad” y “puntualidad” describen cómo de bien representan los datos de la situación actual del mundo real. Cuando se mide la actualidad (lo bien que los datos representan el mundo real de hoy), el resultado solo será válido durante un corto periodo de tiempo. Por ejemplo, después de un año, el resultado almacenado de una posible medida de actualidad será incorrecto, es decir, no nos dirá lo bien que el conjunto de datos representa la situación del mundo real de hoy, sino lo bien que lo representaba hace un año.

La completitud y la exactitud describen ambas lo bien que el conjunto de datos representa el mundo real en un momento determinado.

F.2.3 Orden en la evaluación de la calidad de los datos

Cuando se evalúan datos geográficos, un error individual puede influir en varios elementos de la calidad de los datos. Para las mediciones expresadas en índices (por ejemplo, índices de porcentaje de aspectos de la completitud) es importante el uso de denominadores apropiados para describir la población total; véase la figura F.1.

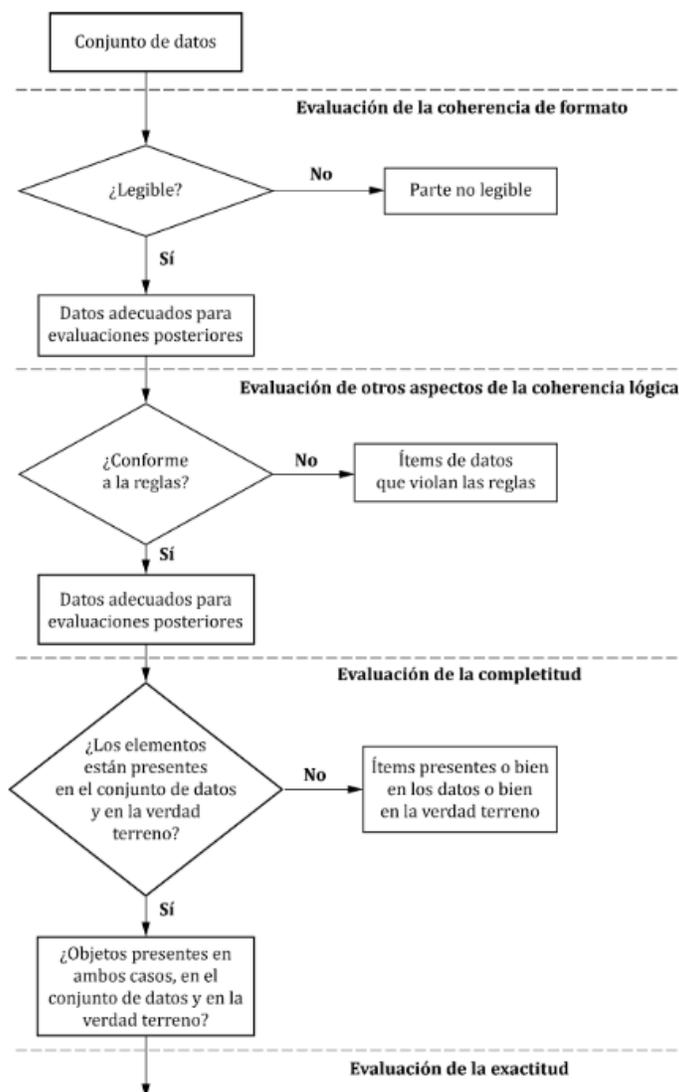


Figura F.1 – Orden en la evaluación de la calidad de los datos

Cuando se evalúa la calidad de los datos, el orden usual es el siguiente:

- a) Coherencia lógica/coherencia de formato: Lo primero que tiene que evaluarse es la legibilidad (o interpretabilidad) de los datos para decidir si los datos se pueden o no decodificar/leer/entender. Se debería informar de los datos no interpretables e ignorarlos en las evaluaciones posteriores. El resultado de la coherencia de formato debería describir qué partes del conjunto de datos no son legibles.
- b) Coherencia lógica: Decidir si se siguen las reglas establecidas para el conjunto de datos. Las partes del conjunto de datos no conformes a las reglas deberían ignorarse en evaluaciones posteriores.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- c) **Completitud:** El siguiente paso en la evaluación es al existencia de los objetos geográficos, aspecto cubierto por la completitud. Para evaluar esto, se comparan los objetos geográficos del conjunto de datos y los datos de la verdad terreno, y se informa de comisiones y omisiones. En la mayoría de los casos, esto no es posible para el conjunto de datos completo. Se someterá a ensayo una muestra del conjunto de datos de acuerdo con el anexo E.
- d) **Exactitud (aspectos posicionales, temático y temporal):** El último paso de la evaluación cubre el aspecto de exactitud, midiendo la diferencia entre las propiedades del objeto geográfico registradas y la verdad terreno. Estas medidas pueden basarse solo en partes del conjunto de datos presentes tanto en el conjunto de datos como en el universo.

F.3 Las relaciones entre los elementos de la calidad de los datos

F.3.1 Generalidades

Muchos elementos de calidad de los datos están relacionados entre sí. En algunos casos esto puede conducir a incertidumbre sobre cómo debería informarse sobre las desviaciones/errores identificados en los datos. Este capítulo discute las relaciones entre los elementos de calidad de los datos.

F.3.2 Elementos de calidad de los datos relativos a valores de atributos omitidos

Al menos tres valores distintos deberían considerarse para indicar que “no hay valor disponible”. El modo en el que usan estos tres puede influir en el elemento de calidad de los datos elegido para informar sobre el valor omitido. Los tres valores tienen significado diferente:

- El valor vacío: en este caso el atributo no tiene ningún valor.
- El valor no aplicable: indica que para este objeto geográfico específico el atributo no es válido, es decir, no tiene significado.

EJEMPLO 1: Fecha de muerte de personas vivas.

- El valor desconocido: en este caso, el atributo válido, es decir, debería haber habido un valor, pero el valor no se conoce.

Los atributos obligatorios con valores vacíos deberían reportarse como errores de coherencia lógica. Los atributos obligatorios no aplicables no deberían contarse al evaluar la completitud de atributo. La cantidad de ocurrencias de valor desconocido debería reportarse como completitud del atributo.

Una forma de aumentar la completitud de atributo es añadir valores artificiales al conjunto de datos. Haciendo esto, el conjunto de datos llegará a ser mejor desde un punto de vista de la coherencia de atributo, pero la exactitud del atributo disminuirá.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

EJEMPLO 2: Un conjunto de datos tiene 50 instancias de objeto geográfico del tipo Árbol. 45 de ellos tienen un valor de atributo almacenado para el atributo HeightOfTree. La exactitud de este atributo (45 instancias) se estima en ± 1 m (desviación típica) y la completitud de atributo es 45/50, es decir 90%. Sin embargo, si a estos valores omitidos de HeightOfTree se les hubiera dado un valor incorrecto (ficticio), por ejemplo, de 10 m, entonces la completitud de atributo llegaría a ser mejor (100%), pero la exactitud de atributo probablemente sería peor.

F.3.3 Relaciones entre los diferentes aspectos de la exactitud

Las discrepancias entre los datos reales y el universo de discurso se pueden medir usando la exactitud posicional, la exactitud temporal y la exactitud temática. Ejemplos de maneras alternativas de expresar la discrepancia son:

- Atributo frente a espacio: Para atributos en los que la distribución geográfica es conocida, la discrepancia puede expresarse tanto por la componente posicional como por la componente temática. El valor de altura de una curva de nivel puede considerarse como un atributo de la curva de nivel. La discrepancia entre la posición actual y la posición real puede medirse tanto por la componente de atributo (“medio metro más alto”) o por componente espacial (“la curva de nivel tiene un sesgo de 10 m en dirección norte”).
- Espacio frente a tiempo: Si se conoce el movimiento de un objeto geográfico, la diferencia entre la posición medida y la posición real puede expresarse tanto por la componente temporal como por la componente posicional, por ejemplo, el error posicional de un coche moviéndose a lo largo de una carretera puede expresarse tanto como “la posición dado podría haber sido correcta hace 20 s” o “la posición es ahora 400 m más allá”).
- Atributo frente a tiempo: “El precio ($\$/m^2$) para una parcela concreta es incorrecto en \$20” o “este era el precio correcto hace 10 años”.

F.3.4 Dependencia entre completitud y exactitud

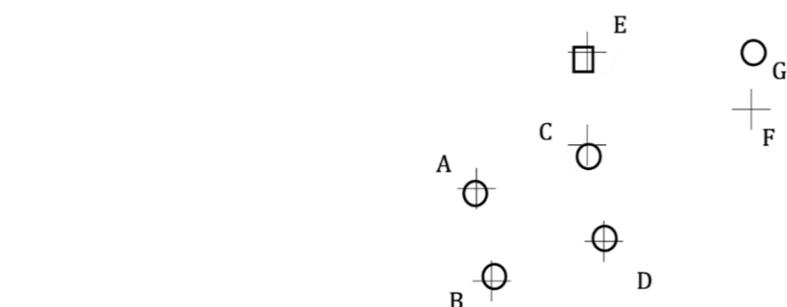
Usualmente, la evaluación de la completitud se basa en comparar el conjunto de datos y el universo de discurso. La operación crítica es el enlace entre los objetos geográficos del conjunto de datos con los del universo de discurso. Cuando existe un identificador único, el enlace se basa usualmente en él.

- a) La cercanía temática (usualmente expresada como tipo de objeto geográfico);
- b) La cercanía geográfica de los objetos geográficos.

Cuando se decide que dos objetos geográficos (un par con un elemento en el conjunto de datos y otro en la verdad terreno) sean representaciones del mismo fenómeno del mundo real, las discrepancias entre los dos se manejan como exactitud. Si se decide que

los objetos geográficos representan objetos distintos, la discrepancia entre los dos se informa usando la completitud (omisión y/o comisión).

Por ejemplo, al evaluar la completitud y exactitud para el tipo de objeto geográfico 1 (véase la figura F.2), no hay problema en las posiciones A, B, C y D. Aquí la clasificación es idéntica (discrepancia temática igual a 0) y las discrepancias geográficas entre las posiciones existentes y reales están dentro del nivel aceptado. En la posición E, las dos instancias tienen una clasificación temática distinta, aunque posicionalmente están muy cerca una de otra. Se tiene que tomar una decisión sobre la discrepancia en la clasificación está dentro del nivel de aceptación para el enlace. Si la respuesta es sí, las dos instancias contribuirán a la evaluación de la exactitud (posicional y/o temática), si la respuesta es no, estamos frente a un caso de completitud (un punto omitido y otro en exceso). En las posiciones F y G, las dos instancias tienen la misma clasificación, pero difieren en posición. Si la discrepancia geográfica se considera dentro del nivel de aceptación para el enlace, la discrepancia contribuirá a la exactitud posicional (probablemente un atípico); si la respuesta es no, estamos frente a un caso de completitud (omisión y comisión).



Leyenda

- Objetos geográficos clasificados como tipo 1 (decisión tomada)**
- + Posición verdadera en el terreno
- O Posición del conjunto de datos real
- Objetos geográficos clasificados como tipo 1 (decisión aún por tomar)**
- Posicionamiento del conjunto de datos real

Figura F.2 – Exactitud frente a la completitud

F.4 Elementos de calidad de los datos. Ejemplo de uso. Ejemplo de uso.

F.4.1 Completitud

F.4.1.1 Generalidades

La presencia y ausencia de objetos geográficos puede describirse por los elementos comisión y omisión de calidad de los datos. Principalmente, la completitud debería utilizarse a nivel de tipo de objeto geográfico, describiendo si los objetos geográficos del universo de discurso se encuentran, o no, en el conjunto de datos.

La completitud también puede ser relevante para propiedades del objeto geográfico (“completitud de atributo” y “completitud de relaciones”). Antes de usar la completitud para esto, deberían considerarse minuciosamente la coherencia lógica y la coherencia conceptual.

F.4.1.2 Comisión – datos presentes en exceso en un conjunto de datos

Puede aplicarse a nivel de instancia de objeto geográfico. Esto significa que los datos se consideran “en exceso” si se trata de toda la instancia de objeto geográfico. Si hay datos no requeridos dentro de una instancia de objeto geográfico o atributo de una instancia de objeto geográfico, entonces no se consideran comisiones.

Esta definición incorpora instancias de objeto geográfico que están presentes en el conjunto de datos pero que no están dentro del ámbito (como se define en las especificaciones).

La regla para los ejemplos siguientes se define como: “solo los objetos geográficos presentes en el universo de discurso deben incluirse en el conjunto de datos”.

EJEMPLO 1: Presencia de datos de Escocia cuando están excluidos del ámbito del conjunto de datos (Inglaterra).

EJEMPLO 2: Solo se pretende incluir en el conjunto de datos de los edificios que son mayores de 5 m² se reporta como comisión.

F.4.1.3 Omisión – datos ausentes de un conjunto de datos

Del mismo modo que para la comisión, puede aplicarse al nivel de instancia de objeto geográfico. En la práctica se refiere a la ausencia de instancias de objeto geográfico cuya inclusión es especificada en las especificaciones.

La omisión debería usarse principalmente cuando falta “todo el ítem” (por ejemplo, una instancia de objeto geográfico). Si se ha omitido una parte obligatoria del ítem (por ejemplo, un atributo obligatorio de una instancia de objeto geográfico), debería informarse como un error de coherencia conceptual.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

La regla para el siguiente ejemplo se define como: “Todas las propiedades residenciales en Inglaterra y Gales deben incluirse en el conjunto de datos”.

EJEMPLO: Ausencia de una propiedad residencial en Inglaterra o Gales en el conjunto de datos.

F.4.2 Coherencia lógica

F.4.2.1 Generalidades

El grado de adherencia a las reglas lógicas de la estructura de datos, la atribución y relaciones (la estructura de datos puede ser conceptual, lógica o física) puede describirse mediante los siguientes elementos de calidad de los datos.

F.4.2.2 Coherencia conceptual – adherencia a las reglas del modelo conceptual

Usualmente, las aplicaciones tienen un esquema conceptual que describe los requisitos de la estructura de datos. Este esquema conceptual puede incluir:

- El nombre de todas las clases (tipos de objeto geográfico, tipos de datos, etc.);
- El nombre de los atributos para todas las clases y también las limitaciones de multiplicidad;
- Los dominios para todos los atributos;
- Las relaciones entre las clases;
- Las relaciones topológicas entre tipos de objeto geográfico, por ejemplo, la relación entre un área y las líneas de borde;
- La relación entre atributos de tipo de objeto geográfico para diferentes tipos de objeto geográfico, por ejemplo, la relación entre el valor de altura sobre el nivel de mar de una curva de nivel y el mismo para una carretera en el punto de cruce geográfico para las dos instancias de objeto geográfico.

F.4.2.3 Coherencia de dominio – adherencia de valores a su dominio

Usualmente, los dominios de valores están descritos por el esquema conceptual de la aplicación y pueden reportarse como parte de la coherencia conceptual o como coherencia de dominio. Si en el esquema conceptual las definiciones de dominio no existen o no son válidas, entonces solo podrá usarse el elemento de calidad coherencia del dominio.

EJEMPLO 1: Una organización define los dominios de valor válidos para cada campo en términos de longitud, tipo de datos y contenido. La coherencia de dominio se usa para asegurar la conformidad con estas condiciones con las siguientes excepciones:

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Cuando el campo contiene datos de posición (es decir, coordenadas latitud y longitud), en cuyo caso se considera como exactitud posicional;
- Cuando el campo contiene datos de fecha/tiempo, en cuyo caso se considera como calidad temporal;
- Cuando el campo contiene una clave primaria, en cuyo caso se considera como coherencia lógica.

La regla para el siguiente ejemplo se define como: El campo LENGUAJE debe contener “ENG” o “CYM”.

EJEMPLO 2: Ejemplo de error de coherencia de dominio: “COR”.

F.4.2.4 Coherencia de formato – grado en que los datos se almacenan de acuerdo con la estructura física del conjunto de datos

La coherencia de formato debería usarse principalmente como primera prueba de evaluación de la calidad para determinar si el conjunto de datos tiene el formato correcto de acuerdo con las especificaciones (de producto).

Si se definen ciertas reglas para establecer el formato de atributos específicos (por ejemplo, para ID generados) la coherencia de formato puede ser relevante para valores únicos de atributo. Si los valores de atributo se comprueban por comparación con lista de valores legales (un dominio), se debería utilizar la coherencia de dominio.

EJEMPLO 1: Las especificaciones de producto de datos de un producto establece el lenguaje geográfico de marcado (GML) como formato de distribución. Si el conjunto de datos no es un archivo GML, entonces este error debería consignarse como error de coherencia de formato. Si un único ítem en el archivo GML está “en formato incorrecto”, por ejemplo, texto en vez de número, esto puede informarse como error de coherencia conceptual o error de coherencia de dominio.

EJEMPLO 2: En un conjunto de datos de una red con requisitos vagos en el esquema conceptual para una “red limpia”, las “partes sucias” (faltas, excesos, superposiciones, autointersecciones, etc.) se consignan como errores de coherencia topológica.

F.4.3 Exactitud posicional

La exactitud de la posición de los objetos geográficos en relación con la Tierra puede describirse utilizando los elementos de calidad de los datos de este apartado.

Medir la exactitud posicional usando verdad terreno implica establecer “correspondencia de pares” con una instancia de objeto geográfico del conjunto de datos y su correspondiente en el conjunto de datos de control (verdad terreno). Si los objetos geográficos tienen identificadores únicos (por ejemplo, respecto a parcelas catastrales), la correspondencia se puede establecer usando identificadores. Los errores groseros,

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

los sesgos y las desviaciones típicas pueden estimarse y reportarse como exactitud posicional.

Sin identificadores disponibles, la correspondencia tiene que establecerse usando las posiciones. Debe definirse un “límite de distancia de correspondencia”. Esto hace que sea imposible calcular los errores groseros. Este “límite de distancia de correspondencia” debe documentarse en el informe. En este caso:

- Las instancias de objeto geográfico en el conjunto de datos sin su correspondiente instancia de objeto geográfico en el conjunto de datos de control deberían informarse como completitud/comisión;
- Las instancias de objeto geográfico del conjunto de datos de control sin su correspondiente instancia de objeto geográfico en el conjunto de datos deberían informarse como completitud/omisión.

F.4.4 Calidad temporal

F.4.4.1 Generalidades

La exactitud de los atributos temporales y de relaciones temporales de los objetos geográficos puede describirse utilizando los siguientes parámetros de calidad de los datos.

F.4.4.2 Exactitud de una medida de tiempo – proximidad a valores verdaderos o aceptados como verdaderos de las medidas de tiempo reportadas

EJEMPLO: Dentro de una cierta organización la exactitud una medición de tiempo se usa para asegurar que:

- El valor no contraviene una condición específica impuesta para el equipo (por encima o debajo las condiciones impuestas por la naturaleza del dato de fecha/hora).

Ejemplo de regla: El campo START_DATE no puede contener un valor futuro.

F.4.4.3 Coherencia Temporal – corrección del orden de los eventos

Las reglas que describen la “corrección del orden de los eventos” puede ser parte del esquema conceptual. Esto puede informarse tanto como coherencia temporal o como coherencia conceptual si las reglas son parte del esquema conceptual.

EJEMPLO: Dentro de una cierta organización, la coherencia temporal se usa para:

- Confirmar la coherencia entre valores de fecha/hora en la relación con el ciclo de vida del objeto del mundo real;

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Asegurar la coherencia de valores de fecha/hora usados en la gestión de las instancias de objeto geográfico en el conjunto de datos.

Ejemplo de regla: La fecha END_DATE debe ser igual o posterior a la fecha START_DATE.

Ejemplo de error de coherencia temporal: START_DATE = “2020-02-02” y END_DATE = “2000-01-01”.

F.4.4.4 Validez temporal – validez de datos con respecto al tiempo

Las reglas que describen “la validez de datos con respecto al tiempo” pueden ser parte del esquema conceptual. Puede informarse tanto de validez temporal como de coherencia conceptual si las reglas forman parte del esquema conceptual.

EJEMPLO: Dentro de una cierta organización la exactitud de una medida del tiempo se utiliza para:

- Asegurar que el contenido de un campo de fecha u hora tiene el formato correcto y usa el calendario definido en las especificaciones.

Ejemplo de regla: El valor de la fecha debe estar en formato ISO 8601-1 – “CCYY-MM-DD”.

Ejemplo de error de validez temporal: “01-01-2010” o “2010-51-15”.

F.4.5 Calidad temática

F.4.5.1 Generalidades

La exactitud de atributos cuantitativos y la corrección de atributos no cuantitativos y de la clasificación de objetos geográficos y sus relaciones pueden describirse usando los siguientes elementos de calidad de los datos.

F.4.5.2 Corrección de la clasificación – comparación de las clases asignadas a objetos geográficos o a sus atributos frente a un universo de discurso (por ejemplo, verdad terreno o conjunto de datos de referencia)

Dentro de una cierta organización, esta definición se usa de manera muy estricta. Clasificaciones que no están definidas dentro de las especificaciones del conjunto de datos no se consideran en la evaluación de la corrección de la clasificación (se consideran en la evaluación de la coherencia de dominio).

Por lo tanto, cuando se evalúa la completitud a nivel de objeto geográfico es recomendable ser consciente de que algunos errores de comisión u omisión pueden provenir de problemas de clasificación errónea. Esto puede ser útil para ofrecer

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

información sobre corrección de la clasificación, pero entonces el error se informará dos veces.

Para evitar reportar dos veces los errores, es posible informar sobre completitud a un nivel superior (conjunto de datos, agrupamiento de tipos de objeto geográfico, etc.) y sobre la clasificación errónea a nivel de objeto geográfico.

EJEMPLO: Se requiere que todas las viviendas tengan una dirección. Este requisito no puede evaluarse correctamente si algunas viviendas están clasificadas como almacén o garaje.

Se ofrece otro ejemplo similar en el anexo D.

F.5.2 Elementos de calidad relacionados con identificadores únicos

A continuación, se presentan algunos casos de uso asociados a elementos de calidad de los datos relevantes para cuestiones relativas a identificadores únicos, véase la tabla F.1.

Tabla F.1 – Elementos de calidad relacionados con identificadores únicos

Casos de uso	Elemento de calidad de los datos que se considera
Todos los identificadores deben tener un formato que se adecúe a las reglas que los definen.	Coherencia de dominio y coherencia de formato.
Todos los identificadores utilizados son válidos de acuerdo con una lista de identificadores únicos reservados.	Coherencia de dominio.
La misma instancia de objeto geográfico está presente dos veces con un mismo identificador único.	Coherencia conceptual de la completitud (los identificadores únicos deben ser únicos).
La misma instancia de objeto geográfico está presente dos veces con diferentes identificadores únicos. NOTA: Aquí el reto es asegurar que las dos instancias de objeto geográfico son realmente dos representaciones del mismo objeto del mundo real.	Comisión

**ANEXO G
(informativo)
Agregación de resultados de la calidad**

G.1 Introducción

Usualmente, una evaluación basada en un único elemento de calidad de los datos no es suficiente para satisfacer al usuario. Usualmente, el productor de datos establecerá (es de esperar que en cooperación con los usuarios potenciales del producto) unas especificaciones de producto de datos ofreciendo todos los requisitos establecidos para el producto.

Para un usuario potencial, será una gran ventaja encontrar una declaración que describa que el producto está evaluado tomando como base unas especificaciones. Tal declaración es un resultado agregado de la calidad de los datos y también puede ser útil en otras situaciones diferentes a la de informar de la conformidad con unas especificaciones.

La calidad de un conjunto de datos puede representarse por uno o más resultados agregados de la calidad de los datos (RACD). El RACD combina los resultados de la calidad a partir de evaluaciones de la calidad de los datos basadas en diferentes elementos de la calidad de los datos o diferentes ámbitos de la calidad de los datos.

Ejemplos de métodos que pueden utilizarse para producir un RACD se indican en los capítulos G.2 a G.4. Se puede considerar que un conjunto de datos es de una calidad agregada aceptable a pesar de que uno o más de los resultados de la calidad de los datos individuales falle en la aceptación. Por tanto, la agregación debería utilizarse solo cuando existen razones de peso. El significado del resultado agregado de calidad de los datos debería quedar siempre claro.

Dado que puede ser difícil de entender completamente un RACD, su significado debería comprenderse antes de extraer conclusiones para el conjunto de datos basadas en él.

El apartado 11.2.1 describe cómo reportar resultados agregados de la calidad de los datos.

G.2 Apto/no apto al 100%

A cada resultado de la calidad implicado en el cálculo se le da un valor booleano de uno (1) si ha pasado la prueba y de cero (0) si ha fallado. La calidad agregada, Q_{RACD} , se determina mediante la fórmula (G.1):

$$Q_{RACD} = V_1 * V_2 * V_3 \quad (G.1)$$

Donde n es el número de marcos de medida de la calidad.

Si $RACD = 1$, entonces la calidad global del conjunto de datos se considera totalmente conforme, por lo tanto, es apto. Si $RACD = 0$, entonces se considera no conforme, por lo tanto, no es apto. Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

tanto, falla. Esta técnica no ofrece un resultado que indique la localización o magnitud de la no conformidad.

G.3 Apto/no apto ponderado

A cada resultado de la calidad implicado en el cálculo se le asigna un valor Booleano de uno (1) si ha pasado la prueba y de cero (0) si ha fallado. Basándose en su importancia para el propósito del producto, a cada resultado de la calidad de los datos se le asigna un peso entre 0 y 1, ambos inclusive. La suma de todos los pesos, w , debería ser igual a 1. La elección de pesos es una decisión subjetiva realizada por el productor de datos o por el usuario. La razón de la decisión del productor de datos debería reportarse como parte del resultado. La calidad agregada, Q_{RACD} , se determina mediante la fórmula (G.2):

$$Q_{RACD} = V_1 * W_1 + V_2 * W_2 + V_3 * W_3 + \dots + V_N * W_N \quad (G.2)$$

Donde n es el número de marcos de medida de la calidad.

Esta técnica proporciona un valor cuya magnitud indica lo cerca que está un conjunto de datos de la conformidad plena. La técnica no proporciona un valor cuantitativo que indique dónde ocurre la conformidad o la no conformidad.

EJEMPLO: Se prepara una tabla de error (véase la tabla G.1) para mostrar el número de errores encontrados y cómo se clasifican según un procedimiento típico para bases de datos de carreteras. El procedimiento de este ejemplo particular asigna pesos a cada tipo de error.

Tabla G.1 – Ejemplo de cálculo de un resultado agregado de evaluación de la calidad

Objeto geográfico	Número de ítems en el lote	Número de ítems no conformes	Razón de ítems no conformes	Proporción de exactitud (definida como 1-ratio)	Pesos	Valor ponderado (proporción de exactitud × peso)
Sección de carretera	19					
Incorrecto		1				
Falta		0	4 / 19	0,79	50 %	0,3950
Exceso		3				
Nombre de la calle						
Nombre base	19	5	5 / 19	0,74	15 %	0,1110
Sentido	19	1	1 / 19	0,95	25 %	0,2375
Hidrografía	1	0	0 / 1	1,00	10 %	0,1000
Exactitud total	(definida como la suma de las proporciones de exactitud ponderadas *100)					84,35 %
NOTA 1 Un ítem se define como una sección de carretera cuando está limitado por puntos de intersección con otras carreteras o con límites de la unidad de muestreo.						
NOTA 2 La agregación de información de calidad de los datos, especialmente utilizando pesos, no tiene gran significado para los usuarios finales y puede ser engañosa dependiendo de los pesos utilizados por el productor de datos.						

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

G.4 Valor máximo/mínimo

A cada resultado de calidad de los datos se le asigna un valor, v , basado en su importancia para el propósito del producto. La razón de la decisión del productor de datos debería ser reportada como parte del resultado. La calidad agregada se determina mediante cualquiera de estas ecuaciones, fórmula (G.3) o fórmula (G.4).

$$Q_{RACD} = \text{MÁX.}(v_i, \text{in} = 1 \dots n) \quad (\text{G.3})$$

$$Q_{RACD} = \text{MÍN.}(v_i, \text{in} = 1 \dots n) \quad (\text{G.4})$$

Donde n es el número de marcos de medida de la calidad.

Esta técnica proporciona un valor cuya magnitud indica lo cerca que está un conjunto de datos de la conformidad plena, pero solo en términos del marco de medidas de calidad representado por máximo o mínimo. La técnica proporciona un valor cuantitativo que indica dónde se produce la conformidad o la no conformidad cuando el marco de medidas seleccionado se expresa junto con el RACD. Sin embargo, este tipo de RACD dice poco acerca de la magnitud de los otros resultados de la calidad.

**ANEXO H
(Normativo)
Descripción de la codificación XML**

H.1 Introducción

El esquema de implementación XML de este documento sigue las reglas de la Especificación Técnica ISO/TS 19139-1.

El esquema XML también utiliza los patrones para desvincular espacios de nombres XML descritos en la Especificación Técnica ISO/TS 19115-3:2016, capítulo 8.

NOTA: La Norma ISO 19157:2013 proporcionó una aplicación del esquema XML que ahora se describe en la Especificación Técnica ISO/TS 19157-2. La descripción de la codificación proporcionada en este documento no es una revisión de la Especificación Técnica ISO/TS 19157-2 ni de ninguna aplicación del esquema XML publicada anteriormente. Tras la publicación de este documento, la Especificación Técnica ISO/TS 19157-2:2016 quedará derogada.

H.2 Espacios de nombres XML

Las definiciones de esquema XML pertenecen al siguiente espacio de nombres:

<https://schemas.isotc211.org/19157/-1/dqc/1.0>.

Este espacio de nombre se abrevia como “dqc” (*data quality concepts*), que corresponde a conceptos de calidad de los datos.

Este esquema XML aplica todas las clases UML definidas en este documento e importa todas las clases pertinentes de otras normas internacionales.

H.3 Esquema XML

La ubicación del esquema XML para el espacio de nombres “dqc” es

<https://schemas.isotc211.org/19157/-1/dqc/1.0.0/dqc.xsd>

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

Anexo I
(Informativo)

Compatibilidad con la Norma ISO 19157:2013

El modelo de la calidad de los datos, sus conceptos y medidas conformes a la Norma ISO 19157:2013 no necesariamente son conformes a este documento (Norma ISO 19157-1:2022). Este anexo enumera en detalle los cambios que se han realizado durante la revisión de la Norma ISO 19157:2013.

NOTA: Los encabezados de este anexo corresponden a los encabezados de la Norma ISO 19157:2013.

- 1) Objeto y campo de aplicación
 - o Se han realizado revisiones menores al texto.
- 2) Conformidad
 - o Se han actualizado los requisitos de conformidad.
 - o El capítulo ha pasado a ser el capítulo 5,
- 3) Normas para consulta
 - o Se ha actualizado el listado de normas para consulta.
- 4) Términos y definiciones
 - o Este capítulo se ha revisado sustancialmente:
 - Se han eliminado los siguientes términos: 'catálogo', 'método de evaluación directo', 'método de evaluación indirecto' y 'usabilidad'.
 - El término 'medida básica de la calidad de los datos' se ha trasladado a un nuevo documento (actualmente en desarrollo) sobre el registro de medidas de calidad de los datos (ISO/AWI 19157-3:-²).
 - Se han añadido los siguientes nuevos términos: 'cobertura', 'calidad de los datos', 'medida de la calidad de los datos', 'unidad de calidad de los datos', 'linaje', 'evaluación de la calidad', 'requisito' e 'incertidumbre'.
 - o La terminología de este documento se ha revisado de acuerdo con las normas internacionales vigentes para la calidad de los datos.
- 5) Abreviaturas de términos

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- El listado de abreviaturas de términos se ha actualizado y vuelto a formatear.
- Requisitos generales para la calidad de la información geográfica
 - En este documento se ha definido un nuevo capítulo (capítulo 6) con un resumen de todas las clases de requisito, recomendación y permiso. Siguiendo las pautas del ISO/TC 211 para el diseño modular de normas, se identificaron los requisitos, recomendaciones y permisos que se escribieron directamente en los párrafos de los capítulos de la Norma ISO 19157:2013 y se les asignó un URI que cumple con ISO/TC 211.
- 6) Visión general de la calidad de los datos
 - Se ha actualizado el modelo de la calidad de los datos y el cambio se refleja en las respectivas figuras de las partes del modelo que se han incorporado a lo largo del documento.
 - Se ha suprimido el uso de prefijos de paquetes en el nombre de los tipos.
 - Se ha alineado el documento con la Norma ISO 19131 y la Norma ISO 19115-1.
 - Se ha introducido un conjunto de requisitos y recomendaciones de acuerdo con las mejores prácticas del ISO/TC 211.
 - El capítulo ha pasado a ser el capítulo 7,
- 7) Componentes de la calidad de los datos
 - Se ha suprimido el *elemento usabilidad* del modelo de la calidad (figura 2).
 - Se ha añadido el apartado 8.4 sobre la extensión del modelo de la calidad estándar.
 - El elemento *metacalidad* se ha desplazado al apartado 8.3.7.
 - Se ha revisado la figura 8.
 - El capítulo ha pasado a ser el capítulo 8.
- 8) Medidas de calidad de los datos
 - Se ha añadido el apartado 9.3 sobre las medidas de calidad definidas por el usuario.
 - El capítulo ha pasado a ser el capítulo 9.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- 9) Evaluación de la calidad de los datos
 - o Se ha realizado una revisión menor de la figura 11.
 - o El capítulo ha pasado a ser el capítulo 10.
- 10) Informe de la calidad de los datos
 - o Se ha revisado la figura 13.
 - o El capítulo ha pasado a ser el capítulo 11.
- Se ha añadido un nuevo capítulo 12 que contiene los requisitos para la codificación XML.
 - o Este es un nuevo capítulo que define los requisitos para la codificación XML de la definición de la calidad de los datos.
- Anexo A
 - o Nuevas pruebas de conformidad han sustituido a las pruebas anteriores.
- Anexo B
 - o Se ha revisado la figura B.1
 - o Se han realizado revisiones menores al texto.
- Anexo C
 - o Se han revisado las tablas para que se correspondan los diagramas de los capítulos 7 a 11.
- Anexo D
 - o Se ha trasladado a la Norma ISO 19157-3:--
- Anexo E
 - o Se ha renombrado como anexo D.
- Se han realizado revisiones menos del texto.
 - o Se han eliminado las referencias al 'elemento usabilidad'.
- Anexo F
 - o Se ha renombrado como anexo E.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- Se han realizado revisiones menores del texto.
- Anexo G
 - Se ha trasladado a la Norma ISO 19157-3:-.
- Anexo H
 - Se ha trasladado a la Norma ISO 19157-3:-.
- Anexo I
 - Se ha renombrado como anexo F.
 - Se han realizado revisiones menores del texto.
 - Se han añadido nuevos ejemplos.
- Anexo J
 - Renombrado como anexo G.
- Se ha añadido un nuevo anexo normativo (anexo H en este documento) que detalla la aplicación del esquema XML de este documento.
- Se ha añadido un nuevo anexo informativo (anexo I en este documento) que detalla la compatibilidad de este documento con la versión anterior, la Norma ISO 19157:2013.
- Bibliografía
 - Se ha actualizado la bibliografía.

BIBLIOGRAFÍA

- [1] ISO 2859 (todas las partes), *Sampling procedures for inspection by attributes*.
- [2] ISO 3534-2:2006, *Statistics. Vocabulary and symbols. Part 2: Applied statistics*.
- [3] ISO 3951-1, *Sampling procedures for inspection by variables. Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL*.
- [4] ISO 8000-2:2022, *Data quality. Part 2: Vocabulary*
- [5] ISO 8601-1, *Date and time. Representations for information interchange. Part 1: Basic rules*.

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- [6] ISO 9000:2015, *Quality management systems. Fundamentals and vocabulary.*
- [7] ISO 19101-1, *Geographic information. Reference model. Part 1: Fundamentals.*
- [8] ISO 19103:2015, *Geographic information. Conceptual schema language.*
- [9] ISO 19104, *Geographic information. Terminology.*
- [10] ISO 19105:2022, *Geographic information. Conformance and testing.*
- [11] ISO 19108, *Geographic information. Temporal schema.*
- [12] ISO 19109:2015, *Geographic information. Rules for application schema.*
- [13] ISO 19110:2016, *Geographic information. Methodology for feature cataloguing.*
- [14] ISO 19115-1:2014, *Geographic Information. Metadata. Part 1: Fundamentals.*
- [15] ISO 19116:2019, *Geographic Information. Positioning services.*
- [16] ISO/FDIS 19123-1:-,³⁾ *Geographic Information. Schema for coverage geometry and functions – Part 1: Fundamentals.*
- [17] ISO 19131:2022, *Geographic Information. Positioning services.*
- [18] ISO 19135-1:2015, *Geographic Information. Procedures for item registration. Part 1: Fundamentals.*
- [19] ISO 19156:2011, *Geographic Information. Observations and measurements.*
- [20] ISO 19160-3:2020, *Addressing. Part 3: Address data quality.*
- [21] ISO/IEC 2382, *Information technology. Vocabulary.*
- [22] ISO/IEC Directives, *Part 1, Procedures for the technical work.*
- [23] ISO/IEC Guide 98-3, *Uncertainty of measurement. Part 3: Guide to expression of uncertainty in measurement (GUM:1995)*
- [24] ISO/IEC Guide 99, *International vocabulary of metrology. Basic and general concepts and associated terms (VIM).*
- [25] ISO/IEC 25012:2008, *Software engineering. Software product Quality Requirements and Evaluation (SQuaRE). Data quality model.*
- [26] ISO/IEC 25020, *Systems and software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). Measurement of quality in use.*
- [27] ISO/IEC 25022, *Systems and software engineering. Systems and software*

Este documento comprende un borrador de Proyecto de Norma Técnica desinado exclusivamente para Consulta Pública. No constituye ningún tipo de elemento ni sustento legal. Prohibida su reproducción y o venta.

- quality requirements and evaluation (SQuaRE). Measurement of quality in use.*
- [28] ISO/IEC 25023, *Systems and software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). Measurement of system and software product quality.*
- [29] ISO/IEC 25024, *Systems and software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). Measurement of data quality.*
- [30] ISO/IEC/IEEE 24765:2017, *Systems and software engineering. Vocabulary.*
- [31] ISO/TS 19115-3:2016, *Geographic information. Metadata. Part 3: XML schema implementation for fundamental concepts.*
- [32] ISO/TS 19139-1, *Geographic information. XML schema implementation. Part 1: Encoding rules.*
- [33] ISO/TS 19129, *Geographic information. Imagery, gridded and coverage data framework.*
- [34] LANTMÄTERIET. (2020), *Quality description Lase data*, versión 1.3, [online] Available at: https://www.lantmateriet.se/globalassets/kartor-och-geografisk-information/hojddata/quality_description_lidar.pdf
- [35] LANTMÄTERIET. (2019), *Hydrography download service, product description version 1.1*, [online] Available at: https://www.lantmateriet.se/globalassets/kartor-och-geografisk-information/geodatatjanster/e_pb_hydrografi_nedladdning_v1.1.pdf
- [36] OGC® Geospatial User Feedback Standard: *Concetpual Model*, [online] Available at: <https://docs.ogc.org/is/15-097r1/15-097r1.html>
- [37] ISO 19157-3:-,⁴⁾ *Geographic information. Data quality. Part 3: Data quality measures register.*