

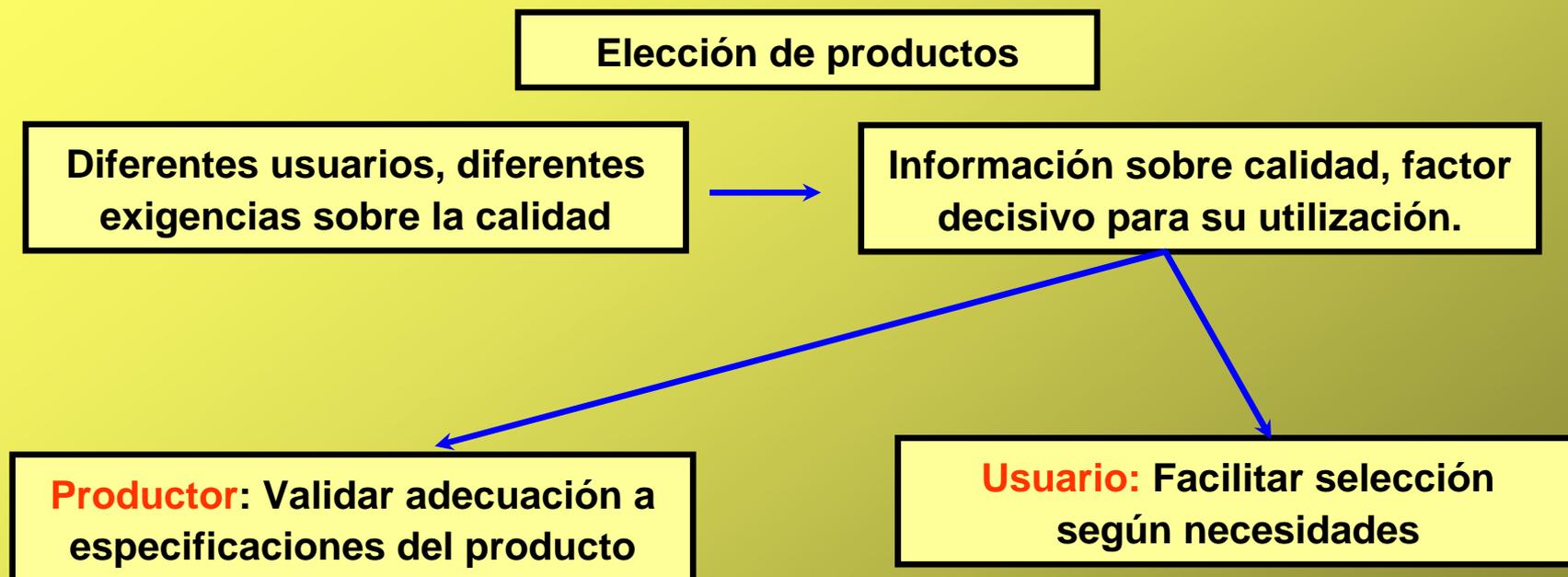
ISO 19138

“Medidas de Calidad de los Datos”



Informar sobre la
calidad significa:

- Identificar** los factores relevantes ISO 19113
- Evaluar** con métodos adecuados ISO 19114
- Cuantificar** utilizando medidas comparables ISO 19138
- Describir** adecuadamente todos los aspectos ISO 19114
ISO 19115



OBJETIVOS:

- Favorecer la buena elección y uso de los productos.
- Incorporar la voz del cliente.
- Competitividad.

El objetivo de esta norma internacional es la normalización de las componentes y estructuras de las medidas de calidad de datos mediante la creación de un registro de medidas de calidad de datos usadas regularmente.



El registro contiene medidas de calidad de los datos que se usan cuando se informa la calidad de los datos para los subelementos de calidad de los datos identificados en la Norma ISO 19113 sobre Principios de Calidad.

El informe deberá realizarse en conformidad con las Normas ISO 19114 - Procedimientos para evaluación de la calidad y 19115 - Metadatos.



La norma proporciona un **grupo inicial de medidas** de calidad de datos basadas en indicadores y **medidas múltiples** definidas para cada subelemento de calidad de los datos.

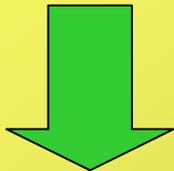
La elección de las medidas a utilizar estará en función del tipo de datos y el propósito perseguido.



Las medidas de calidad de los datos deben estar incorporadas a un **registro** que haya sido elaborado como registro único, en conformidad con la norma ISO 19135 – Procedimientos para el registro de ítems de información geográfica.



La norma ofrece a:



Productores de datos
una guía para informar la
calidad de los datos



Usuarios de datos
una guía para evaluar la utilidad
de un conjunto de datos,
teniendo en cuenta que
diversos usuarios y aplicaciones
frecuentemente tienen
diferentes requerimientos

Esta norma permite a un usuario de datos geográficos comparar la calidad de los conjuntos de datos para determinar cuál satisface los requisitos establecidos.



Para facilitar esas comparaciones, los resultados de los informes de calidad deberán ser expresados de forma comparable.

El uso de medidas incompatibles hace imposible realizar comparaciones en la calidad de los datos.

ISO19138

Es fundamental que la calidad informada para un conjunto de datos contenga **todas** las medidas de calidad que puedan ser de interés para un potencial usuario del conjunto de datos.

Además, es necesario que las mediciones empleadas para determinar la calidad sean informadas y estén **disponibles** al usuario.

Términos y definiciones:

Alias

Abreviatura o nombre corto u otro nombre reconocido, para la misma medida de calidad de los datos.

Corrección

Característica de un ítem que indica que éste no es erróneo.

Medida básica de calidad de los datos

Medida genérica de calidad de los datos usada como base para la creación de medidas de calidad de los datos.

Medida de calidad de los datos

Descriptor cuantitativo de la calidad de datos geográficos, mediante la comparación con su universo del discurso.

Términos y definiciones:

Ámbito de la calidad de los datos

Medida o característica/s de los datos para los que se informa la calidad (ISO 19113).

Tipo de valor de calidad de los datos

Tipo de valor para informar un resultado de la calidad de los datos (ISO 19113).

Error

Cualquier diferencia entre el universo de discurso y el valor obtenido.

Términos y definiciones:

Identificador

Secuencia de caracteres lingüísticamente independiente, capaz de identificar unívoca y permanentemente aquello con lo que está asociado (ISO 19135)

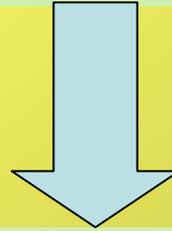
Referencia de la fuente

Referencia a la procedencia de un ítem que ha sido tomado de una fuente externa al registro (ISO 19135).

Universo de discurso

Visión del mundo real o hipotético que incluye todo aquello que es de interés (ISO 19101).

Normas relacionadas:



ISO/IEC Directivas Parte 2 - Reglas para la estructura y el diseño de Normas Internacionales

ISO 19113: 2002 Principios de Calidad

ISO 19114: 2003 Procedimientos para evaluación de la calidad

ISO 19115: 2003 Metadatos

ISO 19131 Especificaciones para productos de datos

ISO 19135 Procedimientos para el registro de ítems de información geográfica

ISO 19139 Metadatos – Especificaciones para su implementación

ISO19138

Objetivos: **Normalizar** las medidas de la calidad del dato.

Cómo?: Definiendo un **conjunto** de medidas normalizadas para la calidad.

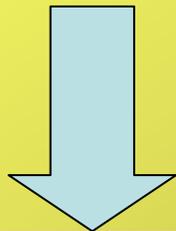
Para qué?: Para **guiar** al productor en la elección de las medidas adecuadas y al usuario en la adecuación de la BDG.

Cuándo usar?: Para **informar** sobre los subelementos cuando se aplican ISO 19113 y 19114

Elementos y subelementos de calidad de los datos, tal como se definen en la Norma ISO 19113.

Elementos de calidad de los datos	Subelementos de calidad de los datos
Compleitud	Comisión
	Omisión
Consistencia Lógica	Consistencia conceptual
	Consistencia de formato
	Consistencia topológica
Exactitud Posicional	Exactitud absoluta o externa
	Exactitud relativa o interna
	Exactitud posicional de los datos en malla
Exactitud Temporal	Exactitud de una medida de tiempo
	Consistencia Temporal
	Validez Temporal
Exactitud Temática	Corrección de clasificación
	Corrección de atributo no cuantitativo
	Exactitud de atributo cuantitativo

Registro



Las medidas de calidad de los datos deberán ser conservadas en un registro.

Dicho registro debe contener componentes técnicos, definiendo las medidas de calidad que deben acompañar a la información administrativa.

El proceso de registración deberá realizarse en conformidad con la norma ISO 19135 - Procedimientos para el registro de ítems de información geográfica.

No todos estos componentes técnicos son obligatorios, los hay optativos y también condicionados al uso previo de algún otro

La descripción para la obligación/condición puede tener los siguientes valores:

-  **M (mandatario)**
-  **C (condicional)**
-  **O (opcional).**



Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Nombre que se da a la medida.
- Si ya existe uno asumido es ese el que se debe usar, en otro caso se propone uno adecuado.

M

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Se refiere a otro nombre o abreviatura reconocidos para referirse a la medida.
- Se puede dar más de un alias.

0

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

- Elemento de la calidad según ISO 19113.

M

Subelemento C. Dato

Medida básica

Definición

Descripción

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Subelemento de la calidad según ISO 19113.

M

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Referencia a la medida básica que es de aplicación en este caso.

C

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Determina el concepto fundamental de la medida.

M

Componentes técnicos

Componentes técnicos

Nombre

Alias

Elemento C. Dato

Subelemento C. Dato

Medida básica

Definición

Descripción

- Describe la medida y sus métodos de cálculo, incluyendo fórmulas, figuras, definición de los tipos de error en los que se basa, etc.

C

Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Variable auxiliar utilizada por la medida **C** de la calidad, se debe incluir nombre, definición, descripción, etc. Puede necesitarse uno o más parámetros. (ejemplo de parámetro es la tolerancia para los overshoots / undershoots)

Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Se refiere al tipo de valor usado para expresar el resultado: booleano, entero, real, etc.

M

Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Se refiere a la estructura que se utiliza para informar sobre el resultado de la calidad del dato. Puede ser: dato único, una serie, una matriz, cobertura, etc.

○

Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Sirve para incluir la cita a la fuente de la medida de la calidad del dato.

C

Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Presenta ejemplos de aplicación de la medida y de su resultado.



Componentes técnicos

Componentes técnicos

Parámetro

Tipo de valor

Estructura

Fuente

Ejemplo

Identificador

- Valor entero que actúa como identificador único de la medida. (ISO 19135).

C

Medidas Básicas

Se introducen para evitar la repetición en la definición de conceptos. Son medidas base para el desarrollo de medidas concretas.

Se relacionan con dos formas de trabajo bien definidas: el conteo de errores y la estimación de la incertidumbre:

Medidas de Conteo:

Se cuentan el número de errores (defectos) o correctos.

Medidas de incertidumbre:

Se basan en modelos estadísticos sobre la incertidumbre de las medidas.

Medidas Básicas

Medidas de conteo		
Indicador de error	Indica que un ítem tiene error	Verdadero (booleano)
Indicador de corrección	Indica que un ítem es correcto	Falso (Booleano)
Casos de error	Indica el número de casos que poseen un error de una tipología dada	5 (Entero)
Casos correctos	Indica el número de casos libres de un error concreto	111 (Entero)
Tasa de error	Número de elementos erróneos frente al total de esos elementos	0.11; 11%; 11:100 (real, porcentaje, ratio)
Tasa de elementos correctos	Número de elementos correctos frente al total de esos elementos	0.21; 21%; 21:100 (real, porcentaje, ratio)

Medidas Básicas

Medidas de incertidumbre (1D)

Probability P	Quantile	Data quality basic measure	Name	Data quality value type
P = 68.3%	$u_{68,3\%} = 1$	$u_{68,3\%} \cdot \sigma_Z$	LE68.3	measure
P = 50%	$u_{50\%} = 0,6745$	$u_{50\%} \cdot \sigma_Z$	LE50	measure
P = 90%	$u_{90\%} = 1,645$	$u_{90\%} \cdot \sigma_Z$	LE90	measure
P = 95%	$u_{95\%} = 1,960$	$u_{95\%} \cdot \sigma_Z$	LE95	measure
P = 99%	$u_{99\%} = 2,576$	$u_{99\%} \cdot \sigma_Z$	LE99	measure
P = 99.8%	$u_{99,8\%} = 3$	$u_{99,8\%} \cdot \sigma_Z$	LE99.8	measure

Medidas Básicas

Medidas de incertidumbre (2D)

Probability P	Data quality basic measure	Name	Data quality value type
$P = 39,4\%$	$\frac{1}{\sqrt{2}} \sqrt{\sigma_x^2 + \sigma_y^2}$	CE39.4	measure
$P = 50\%$	$\frac{1,1774}{\sqrt{2}} \sqrt{\sigma_x^2 + \sigma_y^2}$	CE50	measure
$P = 90\%$	$\frac{2,146}{\sqrt{2}} \sqrt{\sigma_x^2 + \sigma_y^2}$	CE90	measure
$P = 95\%$	$\frac{2,4477}{\sqrt{2}} \sqrt{\sigma_x^2 + \sigma_y^2}$	CE95	measure
$P = 99.8\%$	$\frac{3,5}{\sqrt{2}} \sqrt{\sigma_x^2 + \sigma_y^2}$	CE99.8	measure

Medidas Básicas

Medidas de incertidumbre (3D)

Probability P	Data quality basic measure	Name	Data quality value type
$P = 50\%$	$0.51 \cdot (\sigma_x + \sigma_y + \sigma_z)$	spherical error probable (SEP)	measure
$P = 61\%$	$\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$	Mean radial spherical error (MRSE)	measure
$P = 90\%$	$0.833 \cdot (\sigma_x + \sigma_y + \sigma_z)$	90% spherical accuracy standard	measure
$P = 99\%$	$1.122 \cdot (\sigma_x + \sigma_y + \sigma_z)$	99% spherical accuracy standard	measure

Al igual que en los casos anteriores, esta norma se completa con anexos de gran valor.

Entre estos, se destaca el anexo C dedicado a desarrollar más de 70 medidas de diversa índole, por ejemplo:

- Conteo
- Incertidumbre
- Posición
- Tema
- Coherencia
- etc.

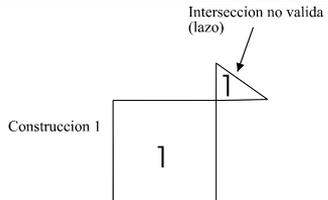


Para cada una de estas medidas se especifican todos los componentes técnicos.

Este listado es interesante, dado que con el amplio abanico que presenta, seguro que se podrá utilizar directamente alguna de las propuestas que contiene.



Ejemplo de medida de conteo aplicada a la consistencia topológica

Línea	Componente	Descripción
1	Nombre	Número de auto-intersecciones erróneas
2	Alias	Lazo
3	Elemento de la calidad del dato	Consistencia lógica
4	Subelemento de la calidad del dato	Consistencia topológica
5	Medida básica de la calidad del dato	Conteo de errores
6	Definición	Conteo de todos los elementos del conjunto de datos que se intersecan ilegalmente consigo mismos.
7	Descripción	--
8	Parámetro	--
9	Tipo de valor	Entero
10	Estructura del valor	--
11	Referencia fuente	--
12	Ejemplo	
13	Identificador	26

Conformidad

Para la **conformidad**, se deben superar todos los requisitos del conjunto de pruebas del **Anexo A**.

Prueba de componentes:

Se verifica que se han incluido todos los componentes de la medida de la calidad de los datos.

Prueba de nombre: Se comprueba si se utiliza un nombre distinto para la medida de la calidad de los datos.

Prueba de elementos/subelementos de la calidad del dato: Se verifica si se asignan un elemento y subelemento de la calidad definidos en ISO 19113 o según ISO 19114 y si la medida de la calidad es relevante para ese elemento.

Prueba sobre medida básica: Se verifica que la medida se deriva de forma apropiada de una medida básica.

Conformidad (cont.)

Para la **conformidad**, se deben superar todos los requisitos del conjunto de pruebas del **Anexo A**.

Prueba de definición: Se verifica si se proporciona una definición correcta, completa y adecuada.

Prueba de descripción: Se comprueba si se da una descripción exhaustiva.

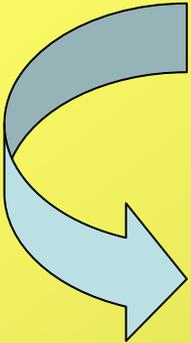
Prueba de parámetro: Se verifica si se incluyen el o los parámetros requeridos.

Prueba tipo de valor : Se determina si se proporciona un tipo de valor adecuado a la medida propuesta.

Prueba de fuente: Se determina si se proporciona una fuente adecuada.

Prueba de ejemplos : Si se incluyen ejemplos, se determina si éstos son válidos para la medida propuesta.

Conclusiones ISO 19113 - 19114 - 19138

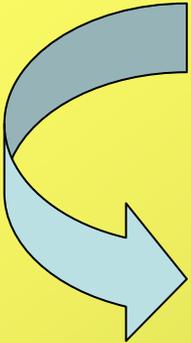


Se dispone de un conjunto de tres normas relativas a la calidad para la información geográfica que se centran en los aspectos complementarios y relativos a:

- **identificar factores relevantes,**
- **evaluar,**
- **medir e**
- **informar.**



Conclusiones ISO 19113 - 19114 - 19138



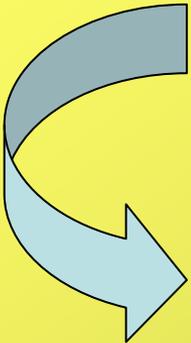
Son tres normas abstractas y generales, que dejan algunos aspectos de la información geográfica sin tratar (por ejemplo: calidad de las imágenes, interoperabilidad de datos, etc.)

Pero permiten que cada quien las amplíe según sus necesidades.

Estas normas no establecen niveles de calidad.



Conclusiones ISO 19113 - 19114 - 19138

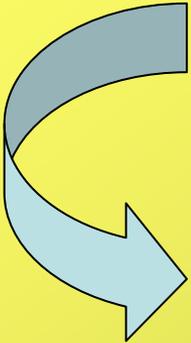


Desde el punto de vista formal son tres documentos con una estructura similar y que pueden interactuar.

Todas éstas normas incluyen anexos de gran interés práctico para el auto-aprendizaje y una lista de verificación de su cumplimiento.



Conclusiones ISO 19113 - 19114 - 19138

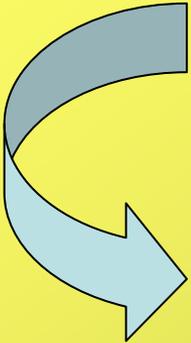


Son normas que se integran en una familia con la que pueden existir algunos problemas de coherencia (por ejemplo ISO 19115).

Su aplicación tiene sentido pleno dentro de la familia pero sus conceptos pueden ser aprovechados para usos fuera de la misma.



Conclusiones ISO 19113 - 19114 - 19138

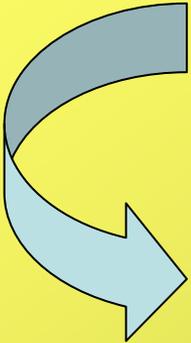


Existe una relación directa con el conjunto de normas ISO dedicadas a la calidad

(por ejemplo: ISO 2859-1, 2859-2, 3951, etc.).



Conclusiones ISO 19113 - 19114 - 19138

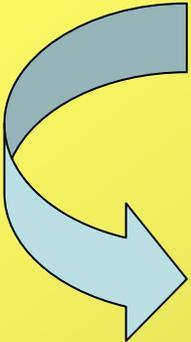


La aplicación de las mismas no es compleja.

La mayor problemática está en el estado de las organizaciones que puedan desear aplicarlas con respecto a la gestión que hacen de la calidad sus productos.



Conclusiones ISO 19113 - 19114 - 19138



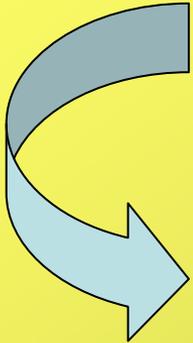
De las tres normas la que puede suponer un mayor esfuerzo de aplicación es la ISO 19113 dado que supone decidir qué es relevante para la calidad.

Esto conlleva conocer muy bien el producto y sus usos.

Lo “relevante” debe estar directamente ligado a la voz del cliente, orientando de esta forma la producción y los procesos propios de la calidad (medidas y evaluación).



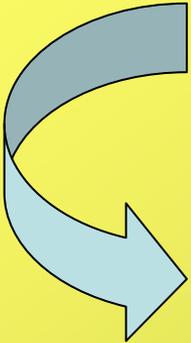
Conclusiones ISO 19113 - 19114 - 19138



ISO 19113 se ha quedado un poco corta en los elementos y subelementos definidos.



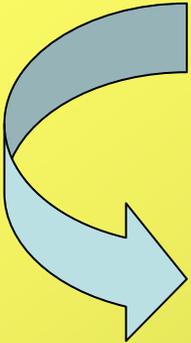
Conclusiones ISO 19113 - 19114 - 19138



**Se centran en la calidad de los datos,
pero los procesos se olvidan.**



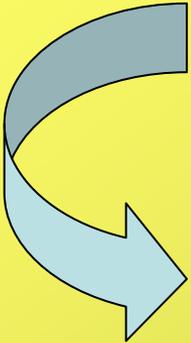
Conclusiones ISO 19113 - 19114 - 19138



No existe una gran experiencia en su aplicación, aunque las instituciones cartográficas más pioneras llevan años utilizando los conceptos base de las mismas.



Conclusiones ISO 19113 - 19114 - 19138



En la actualidad hay un claro interés creciente en los aspectos de la calidad de la información geográfica, lo que está favoreciendo que cada vez sean más los que las aplican.





Paisajes de Catamarca

Gracias por su atención!