

Facultad de Ingeniería

ISO 19157 - Información geográfica - Calidad de datos

Hebenor Bermúdez - Miguel Gavirondo

Octubre 2024



FACULTAD DE
INGENIERÍA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Definiciones

CALIDAD: grado con el que un conjunto de características inherentes cumple unos requisitos. ISO 9000.

CONFORMIDAD: cumplimiento de los requisitos especificados. ISO 19105.

NIVELES DE CONFORMIDAD DE CALIDAD: valor umbral o conjunto de valores umbrales, para los resultados de la calidad de 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 requisitos de usuarios.

UNIVERSO DE DISCURSO: vista del mundo real o hipotético que incluye todo aquello que es de interés. ISO 19101.

¿Por qué medir la calidad?

Ante la **gran abundancia de datos** es necesario contar con herramientas que **permitan su selección** en función de las **necesidades de cada usuario**.

Nuevamente aparecen los actores principales en el proceso: **PRODUCTOR - USUARIO**.

La **confianza en la calidad** es tan importante como el resultado de la calidad. Por lo que hay que contar con **procedimientos claros y comprobados**.

Es necesario que los **resultados sean comparables**. Para esto es necesario que se midan de forma **normalizada** y se informe sobre los resultados de la misma manera.

Objetivos y campo de aplicación

Los objetos de la norma son:

- Define los componentes para describir la calidad de los datos.
- Especifica las componentes y la estructura del contenido de un registro de medidas de la calidad de datos.
- Describe los procedimientos generales para evaluar la calidad de datos geográficos.
- Establece principios para informar sobre la calidad de los datos.

NO ESPECIFICA NIVELES DE CONFORMIDAD

Normas de calidad

- Identificar los factores relevantes. **Sobre que informar.**
- Evaluar con métodos adecuados. **Cómo evaluar cada factor.**
- Cuantificar adecuadamente y de forma comparable. **Qué medidas usar.**
- Describir adecuadamente todos los aspectos. Estructuras, reglas, etc.

PRODUCTOR: tener especificaciones claras para validar los productos contra estas.

USUARIO: Seleccionar los productos y servicios según sus necesidades.

Normas de calidad

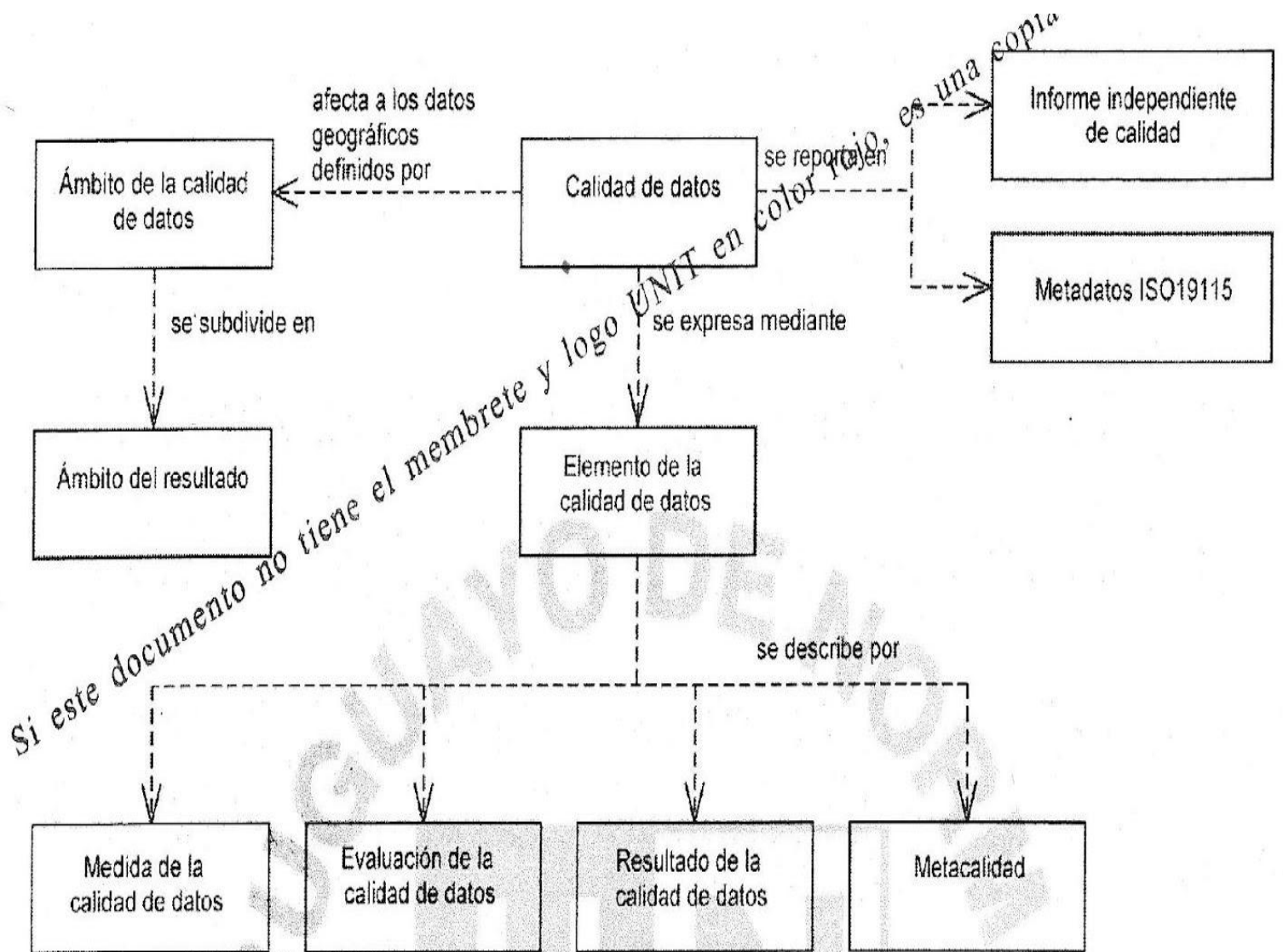


Figura 1 – Modelo conceptual de la calidad para los datos geográficos

Normas de calidad

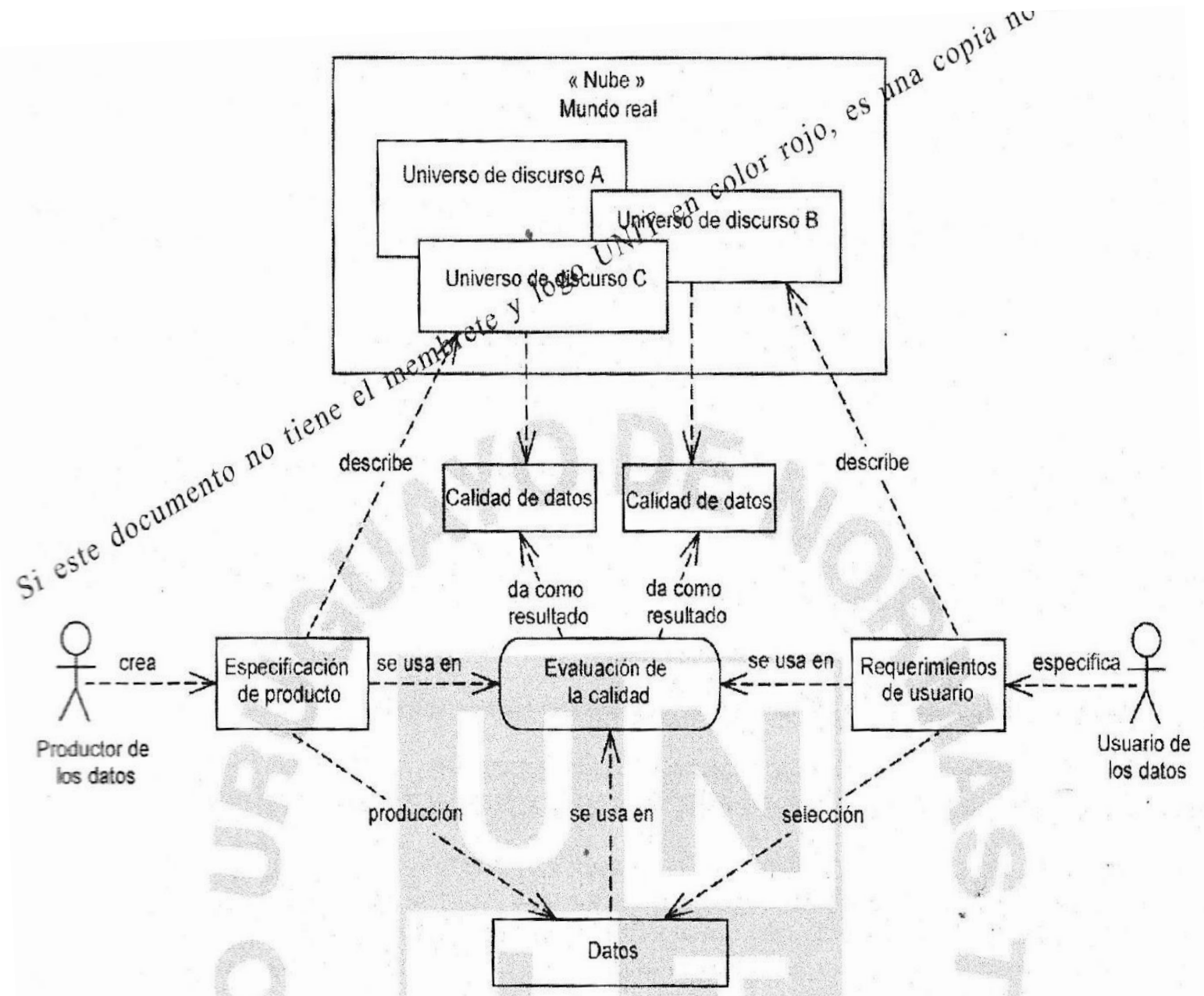


Figura B.1 – Marco de trabajo de los conceptos sobre la calidad de datos

¿Qué evaluamos?

Un conjunto de datos se puede construir para ser usado para una aplicación específica o para ser usado por aplicaciones supuestas.

Lo que se evalúa es **la diferencia entre el conjunto de datos final y el universo de discurso** (el cual debe estar claramente definido en las especificaciones del producto).

El productor va a querer evaluar que tanto el conjunto de datos que produjo **se ajusta al universo de discurso**.

El usuario va a querer evaluar si un conjunto de datos **es adecuado para una aplicación determinada**.

¿Qué evaluamos?

La calidad puede evaluarse de forma **directa** a través de sus elementos de la calidad o de forma **indirecta** a través de información no cuantitativa como el USO, LINAJE o PROPÓSITO los cuales se pueden encontrar en los metadatos.

UNA COSA IMPORTANTE:

Si cambian las especificaciones del producto se debe **validar la evaluación de la calidad** en función de las nuevas especificaciones.

Normas de calidad

Otro objetivo de las normas es **comunicar** a los productores y a los usuarios buscando **evitar** los problemas que surgen de la **visión** que cada actor tiene de la cosa.

Las especificaciones tienen que ser **claras** y **entendibles** a los efectos de que no se generen **ambigüedades** y deben atender a todos los aspectos del objeto.

PUNTOS DE COMUNICACIÓN:

- una característica o aspecto de interés.
- una medida que permita cuantificar esa característica.
- una metodología para obtener la medida.

Elementos de la calidad

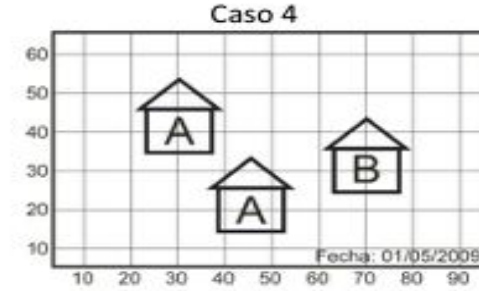
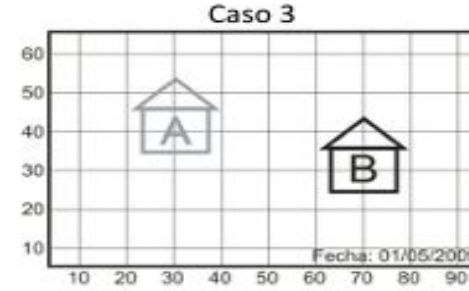
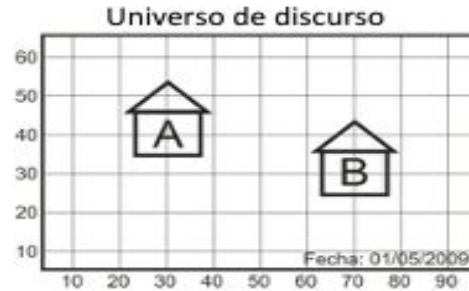
Para cualquier evaluación de la calidad se debe especificar el ámbito sobre el cual se evalúa. El ámbito de la calidad especifica *“la extensión espacial y/o temporal, o la(s) característica(s) común(es) que identifican los datos cuya calidad va a evaluarse”*.

- Una serie de conjuntos de datos.
- Un conjunto de datos.
- Un subconjunto definido por el tipo de objetos, objetos específicos, extensión geográfica, una extensión temporal, etc.

CATEGORÍA COMPLECIÓN: Presencia o ausencia de datos geográficos, sus atributos o sus relaciones

ELEMENTOS: Comisión y Omisión

Categorías y elementos de la calidad COMPLECIÓN

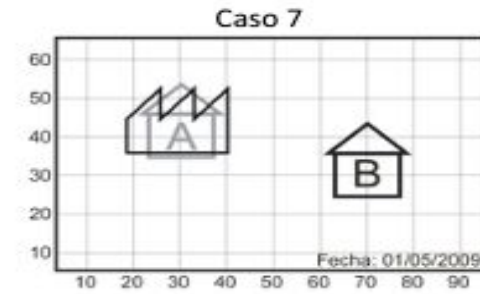
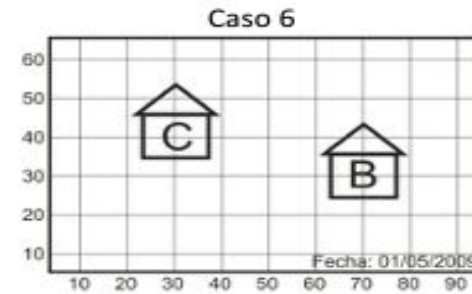


Categorías y elementos de la calidad

CONSISTENCIA LÓGICA

CATEGORÍA CONSISTENCIA LÓGICA: grado de adherencia a las reglas lógicas de la estructura de datos, de los atributos y de las relaciones

ELEMENTOS: Consistencia conceptual, Consistencia de dominio, Consistencia de formato, Consistencia topológica.



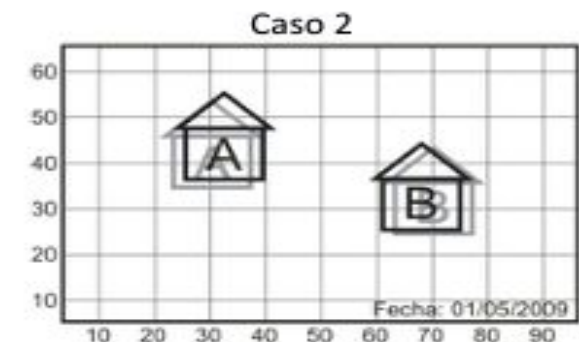
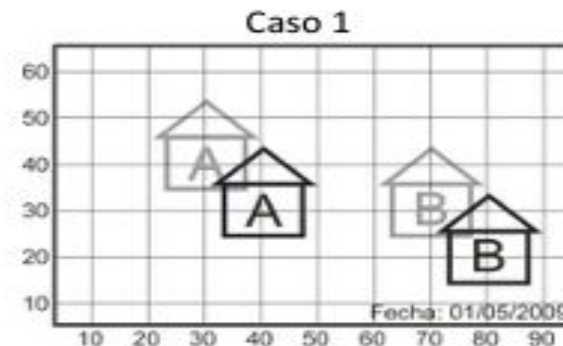
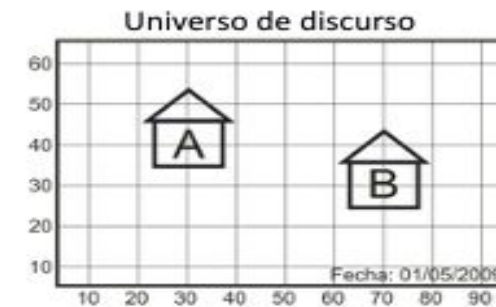
Solo puede contener casas con valores de atributos A o B.

Categorías y elementos de la calidad

EXACTITUD POSICIONAL

CATEGORÍA EXACTITUD POSICIONAL: exactitud de la posición de los objetos geográficos en un determinado sistema de referencia espacial.

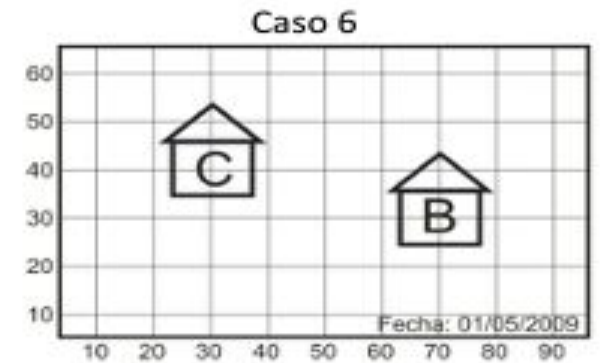
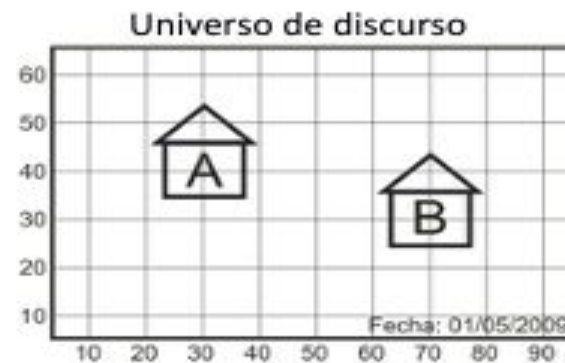
ELEMENTOS: Exactitud absoluta o externa, Exactitud relativa o interna, Exactitud posicional de datos en malla.



Categorías y elementos de la calidad EXACTITUD TEMÁTICA

CATEGORÍA EXACTITUD TEMÁTICA: exactitud de los atributos no cuantitativos y la corrección de las clasificaciones de objetos geográficos y sus relaciones.

ELEMENTOS: Corrección de la clasificación, Corrección de atributos no cuantitativos, Exactitud de atributos cuantitativos.

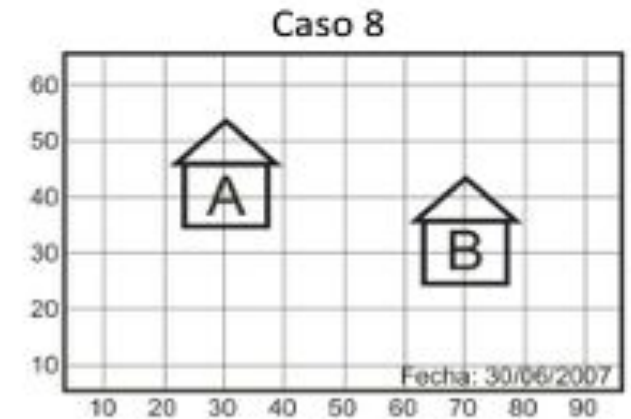
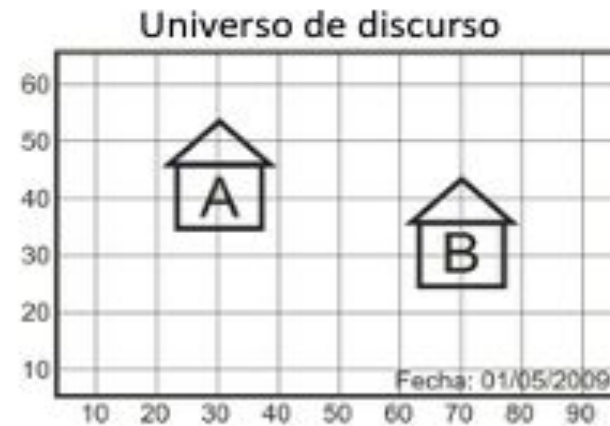


CATEGORÍA EXACTITUD TEMÁTICA: calidad de los atributos y de las relaciones temporales.

ELEMENTOS: Exactitud en una medida del tiempo, Consistencia temporal, Validez temporal.

Categorías y elementos de la calidad

CALIDAD TEMPORAL



Categorías y elementos de la calidad USABILIDAD

ELEMENTO USABILIDAD: este elemento debe usarse para describir información específica de la calidad sobre la idoneidad de un conjunto de datos para una aplicación en particular o sobre su conformidad con un conjunto de requisitos.

Todos los elementos anteriores se pueden usar para evaluar este elemento. Independientemente de esto la usabilidad debe evaluarse cuando alguno de los requerimientos de los usuario no se puede describir con los elementos anteriores.

Descriptores de los elementos

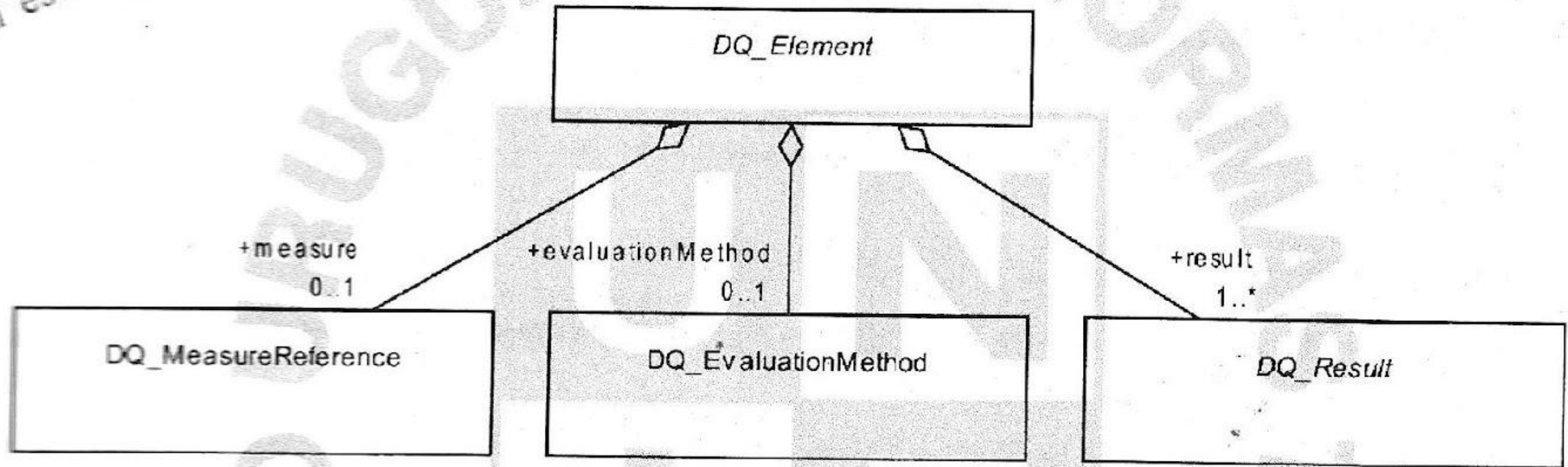


Figura 5 – Descriptores de un elemento de la calidad de datos

Descriptores de los elementos

MEDIDA: cuando se evalúa un elemento se debería hacer referencia a una única medida.

Esto no quita que para **describir correctamente la calidad** de un conjunto de datos se tengan que hacer **varias evaluaciones** (por ejemplo especificando el ámbito) y por lo tanto que se tengan que usar varias medidas. En este caso debe agregarse un “nuevo elemento” evaluado para cada medida.

Descriptores de los elementos

MÉTODO DE EVALUACIÓN: procedimientos y métodos aplicados a los datos geográficos para obtener un resultado de la calidad.

Para cada medida se **debe** proporcionar un método de evaluación.

Cada evaluación **debería** estar acompañada de una fecha o rango de fechas que indique la vigencia de la evaluación.

Descriptores de los elementos

RESULTADO: se debe proporcionar **al menos un resultado** de la calidad de los datos para cada elemento de la calidad.

El resultado puede ser **cuantitativo, de conformidad, descriptivo o de cobertura.**

Pueden brindarse distintos resultados para el mismo elemento y la misma medida.

Descriptorios de los elementos

RESULTADO cuantitativo: puede ser un valor único o un conjunto de valores dependiendo de la medida usada. Este resultado debería incluirse junto con una unidad para el resultado.

RESULTADO de conformidad: este resultado surge de **comparar** el resultado de la evaluación con el valor especificado como nivel de conformidad de la calidad específico y aceptable. Los niveles de conformidad **deben estar indicados** en las especificaciones del producto u otras especificaciones de requisitos. Se pueden especificar distintos resultados de conformidad si se comparan distintos niveles de conformidad.

Descriptores de los elementos

RESULTADO descriptivo: en algunos casos donde no se puede obtener un resultado cuantitativo para la calidad. Para estos casos se puede expresar la calidad a través de una evaluación subjetiva del elemento con una declaración.

RESULTADO de cobertura: es un resultado de la evaluación de la calidad que se organiza como cobertura.

Lectura recomendada

Norma ISO 19157