**Conjuntos de datos y Metadatos**

**¿Qué es un conjunto de datos?**

Un conjunto de datos es una agrupación o colección de datos relacionados, es decir pertenecen a un mismo contexto y están almacenados sistemáticamente para su posterior uso.

El conjunto de datos que puede ser documentado mediante metadatos en un proceso de gestión de información ambiental y sobre biodiversidad, es aquel que cuenta con dos características: i) ser estructurada y ii) ser interoperable. Por estructurada se entiende que los datos o la información tenga un orden lógico que permitan que otro pueda hacer uso de la misma, y por interoperable se entiende a la característica de intercambiar información y posibilitar su puesta en común.

Un ejemplo de un conjunto de datos que se puede documentar son los registros de especies que se encuentran organizados en una tabla de datos de Excel, en una base de datos de Access o cualquier motor de bases de datos. En cuanto a información ambiental documentable están los listados de las áreas protegidas que han sido declaradas en un país o un mapa de las mismas.

**Metadatos**

Una de las definiciones más conocidas de la palabra metadato es “el dato de un dato”. Los metadatos proporcionan la información mínima necesaria para identificar un recurso (Senso y de la Rosa 2003). Estos pueden incluir información descriptiva sobre el contexto, la calidad y la condición o característica del dato de diferentes tipos de recursos.

Los metadatos no solo permiten describir un conjunto de datos, sino también sirven para acceder a los mismos (Cathro 1997). Constituyen un mecanismo para caracterizar datos y servicios de manera que los usuarios puedan encontrarlos y acceder a ellos (Sánchez Maganto et al. 2008).

Un típico ejemplo de lo que puede ser un metadato, es la etiqueta de un producto, la información que se encuentra en ella permite conocer un poco más acerca de este y por ende tomar decisiones acerca de si se desea o no consumir dicho producto de acuerdo a unas prioridades definidas.

Se ha reconocido que los metadatos cumplen funciones como resumir el significado de los datos, permiten su búsqueda, determinar si el dato es el que se necesita, muestra instrucciones de cómo se puede interpretar un dato, da la información de las condiciones de acceso y uso, permite recuperar y usar una copia del dato, también aportan información acerca de la vida del dato y ofrecen información del creador de los mismos (Iannela y Waugh 1997).

Los metadatos se suelen publicar a través de catálogos, los cuales no difieren demasiado de los catálogos tradicionales de una biblioteca, excepto por el hecho de ofrecer una interfaz estandarizada de servicios de búsqueda. Los catálogos son herramientas que permiten la búsqueda, la selección, la localización y el acceso a la información y permiten el contacto entre el productor y el consumidor de la misma (Sánchez Maganto et al. 2008).

Los conjuntos de datos pueden ser insumos a partir de los cuales se generen productos de información como artículos, documentos técnicos, análisis comparativos, mapas, entre otros; esta información **debe ser documentada en el metadato en la sección de referencias cruzadas**. En algunos casos las referencias cruzadas también cuentan con la documentación de su propio metadato, es el caso de los mapas, se debe documentar el metadato de la base de datos que permite generar el mapa y también de la salida gráfica del mapa y el uno es referencia cruzada del otro, pues se complementan.

**Beneficios**

La documentación de metadatos trae consigo un sin número de beneficios, tanto para las instituciones, como para productores y usuarios:

|  |  |
| --- | --- |
| **A las instituciones …**  | * Permite una mejor organización de la información.
* Genera memoria institucional a partir de la documentación de la información producida en el trascurrir de los años. Persiste el conocimiento.
* Permite mantener el valor de la inversión en datos.
* Posibilita el intercambio de datos con otras instituciones.
* Genera ganancias en tiempo y dinero.
* Fortalece los procedimientos de la gestión de información.
* Facilita la ubicación y la detección de los datos.
* Facilita la cooperación interinstitucional en temas relacionados con información.
* Propicia mecanismos para mantener y auditar los datos existentes.
 |
| **A los productores …**  | * Evita duplicaciones en el proceso de producción.
* Reduce esfuerzos de trabajo.
* Permite suministrar información con altos niveles de calidad.
* Permite difundir la información producida.
* Proporciona la información necesaria para describir correctamente sus datos.
* Facilita la organización y el mantenimiento de los metadatos de conjuntos de datos.
 |
| **A los usuarios …**  | * Facilita el descubrimiento, la recuperación y la reutilización de los datos.
* Permite conocer si los datos pueden ser usados para un fin determinado, es decir si son útiles para ellos.
* Brinda información básica sobre los datos disponibles de manera que posibilita su catalogación, así como facilita información de los lugares y formas de almacenamiento de dichos datos.
* Provee información rápida a potenciales distribuidores de datos.
* Facilita la localización, acceso, evaluación, compra y utilización de los datos de manera eficiente.
 |