Propuesta para la estructuración y publicación de datos e información socioeconómica

Beltrán, Néstor
Vélez, Danny
Torres, Marcela

Programa Evaluación y Monitoreo de la Biodiversidad
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt
Bogotá, D.C.
2017

Instituto de Investigación de Recursos Biológicos Alexander von Humboldt
Sede Principal: Calle 28A # 15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2

RESUMEN

Este documento se enmarca en las actividades establecidas para el diagnóstico de los datos y la información socioeconómica generada por el Instituto Humboldt y otras instituciones seleccionadas, el objetivo puntual de este trabajo es realizar la propuesta para estructuración, almacenamiento y publicación de datos e información socioeconómica que sirva como herramienta de gestión dentro del Instituto.

Palabras clave: datos, socioeconómico, diagnóstico, estructuración, gestión, metadato.

ABSTRACT

This document is part of the activities required to the diagnosis of socioeconomic data and information generated by the Humboldt Institute and other select institutions, the main objective of this work is to make a proposal to guide on how to structure, storage and publicate socioeconomic data and information, and to serve as a management tool within the Institute.

Keywords: data, socioeconomic, diagnosis, structuring, management, metadata.

CONTENIDO

RESUMEN 2
ABSTRACT 2
CONTENIDO 3
1. INTRODUCCIÓN 4
1.1. DEFINICIONES 4
1.1.1. Datos socioeconómicos 4
1.2. IMPORTANCIA DE LOS DATOS SOCIOECONÓMICOS 4
1.2.1. Planes de conservación y datos socioeconómicos 4
1.2.2. Información socioeconómica y metas Aichi 4
2. PANORAMA EN OTRAS ENTIDADES 6
2.1. DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA – DANE 6
2.1.1. Proceso de documentación 6
2.1.1.1. Herramienta y estándares utilizados 6
2.1.1.2. Descripción de los elementos necesarios 7
2.1.1.3. Elementos de documentación 7
2.2. INFORMACIÓN SOCIOECONÓMICA EN EL INSTITUTO HUMBOLDT 7
2.2.1. Técnicas utilizadas 7
2.2.2. Acceso 8
2.2.2.1. Tipos de acceso 9
3. PROPUESTA DEL PERFIL 9
3.1. HERRAMIENTA DE PUBLICACIÓN 13
3.1.1. Nesstar Publisher 14
3.1.1.1. Características principales 14
3.1.1.2. Requerimientos 14
3.1.2. Nesstar Server 14
3.1.2.1. Características principales 14
3.1.2.2. Requerimientos 15
5. REFERENCIAS 16

1. INTRODUCCIÓN

1.1. DEFINICIONES

1.1.1. Datos socioeconómicos

Son datos acerca de los humanos, sus actividades y el espacio o estructuras usadas para conducir dichas actividades. Las fuentes para obtener estos datos son diversas: Encuestas de campo, estadísticas, registros administrativos, datos secundarios, etc. Pueden dividirse en datos desagregados: datos acerca de individuos o entidades únicas, por ejemplo: la edad de una persona, sexo, nivel de educación, ingresos, gastos, ocupación, etc., ventas, número de empleados, ganancias, etc. Datos agregados: describen un grupo de observaciones dentro de un criterio definido, por ejemplo: datos geográficos agrupados en unidades espaciales como censos, intervalos de tiempo, socioeconómicos, etc.

1.2. IMPORTANCIA DE LOS DATOS SOCIOECONÓMICOS

1.2.1. Planes de conservación y datos socioeconómicos

La conservación efectiva de la biodiversidad con recursos económicos limitados necesita de acciones que maximicen los resultados obtenidos con el presupuesto asignado. Los resultados o beneficios esperados de las estrategias de conservación dependen siempre de factores biológicos como el número de especies en una región, pero también deben incluir factores socioeconómicos, como la tasa de conversión de tierra en una región. En el escenario de amenazas antropogénicas de gran escala y recursos limitados para mitigar estos efectos, especialmente en los países tropicales de bajos ingresos, se hace necesario fijar las prioridades de conservación. Los datos socioeconómicos permiten estimar eficazmente el nivel de amenaza antropogénica en las regiones con un gran número de especies endémicas (Murdoch et al., 2007).

El objetivo de la conservación es biológico (p.ej. Maximizar el número de especies conservadas) pero los medios para obtenerlo son principalmente socioeconómicos: entender el comportamiento humano que se encuentra en la raíz de las amenazas, diseñar varias estrategias y calcular el costo, crear políticas y demás acciones que puedan dar vuelta al comportamiento humano para que sea más consistente con la conservación. Si no se tienen en cuenta la realidad política, económica y social en la planeación de estrategias de conservación, la implementación de estos puede ser inefectiva (Polasky, 2008).(Polasky, 2008)

1.2.2. Información socioeconómica y metas Aichi

Dentro del plan estratégico para la diversidad biológica 2011-2020 y las metas de Aichi para abordar las causas subyacentes de la pérdida de la diversidad biológica mediante la incorporación de la diversidad biológica en todos los ámbitos gubernamentales de la sociedad, se han resaltado 8 de las metas en las que la información socioeconómica se hace necesaria (Tabla 1).

Tabla 1. Resumen de las metas resaltadas. Se expresa en verde la alta necesidad del tipo de datos e información, en gris la información adicional que puede ser incluida. DGE: datos genéticos, DES: datos de especies, ECO: ecosistemas y servicios ecosistémicos, SEC:datos socioeconómicos, DOT: datos de observación terrestre (Wetzel et al. 2015).

	REQUERIMIENTO DE TIPO DE DATOS				
	Datos de biodiversidad			*Otros datos*	
Metas de Aichi para 2020 Datos e información necesarias	DGE	DES	ECO	**SEC**	DOT
Meta 6 ● Reservas de peces, invertebrados y plantas acuáticas. ● Planes y medidas de recuperación y manejo para estas especies. ● Ecosistemas vulnerables. ● Impactos de la pesca en las reservas, especies y ecosistemas.					
Meta 7 ● Zonas destinadas a la agricultura, acuicultura y silvicultura. ● Conservación de la diversidad biológica.					
Meta 9 ● Especies exóticas invasoras. ● Vías de introducción y establecimiento.					
Meta 10 ● Presiones antropógenas sobre los arrecifes de coral y otros ecosistemas vulnerables.					
Meta 11 ● Sistemas de áreas protegidas y otras medidas de conservación eficaces basadas en áreas.					
Meta 14 ● Ecosistemas que proporcionan servicios esenciales, incluidos servicios relacionados con el agua, y que contribuyen a la salud, los medios de vida y el bienestar. Incluidas las comunidades indígenas y locales.					
Meta 15 ● Resiliencia de los ecosistemas. ● Restauración de las tierras degradadas. ● Contribución de diversidad biológica a las reservas de carbono.					
Meta 18 ● Conocimientos, las innovaciones y las prácticas tradicionales de las comunidades indígenas y locales pertinentes para la conservación y la utilización sostenible de la diversidad biológica.					

2. PANORAMA EN OTRAS ENTIDADES

2.1. DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA – DANE

Adaptado de ANDA - Archivo nacional de datos (2014)

El Archivo Nacional de Datos (ANDA) es un catálogo en el que los usuarios pueden explorar, buscar, comparar, solicitar acceso y descargar información relacionada con censos, encuestas por muestreo y uso estadístico de registros administrativos. El ANDA contiene metadatos de operaciones estadísticas que producen el DANE y otras entidades del Sistema Estadístico Nacional (SEN). Asimismo, algunos de los archivos tienen disponibles microdatos de acceso público. La información publicada en el ANDA está documentada bajo los estándares DDI y Dublin Core (Figura 1).

El proceso de documentación y publicación se resume en tres pasos:

1. Identificación de la información requerida: Aquí se recopilan los documentos que contengan los aspectos metodológicos de la operación estadística a documentar, también la estructura de la base de datos y el material de referencia.
2. Documentación y validación: Se realiza en la herramienta Nesstar Publisher, luego se realiza una revisión y validación de la información para verificar la exactitud y claridad de lo documentado.
3. Difusión: Terminado el proceso de validación se realiza la carga en el Archivo nacional de Datos, para que los usuarios puedan consultarla.

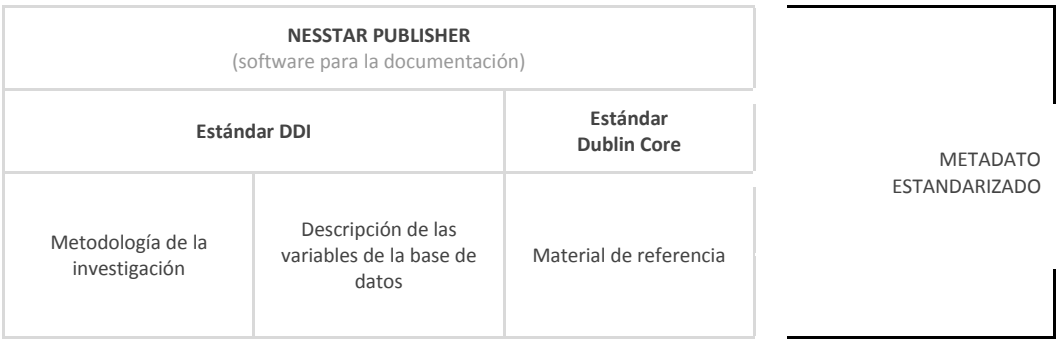

Figura 1. Proceso de producción de un metadato estandarizado. El Nesstar Publisher permite la utilización de los estándares DDI (Descripción de la metodología de la investigación) y Dublin Core (Material de referencia: Manuales, documentos de diseño, etc.).

2.1.1. Proceso de documentación

2.1.1.1. Herramienta y estándares utilizados

Nesstar: Es el editor para la documentación y preparación de los datos y metadatos para su publicación en el catálogo en línea ANDA, está diseñado para promover la adopción de estándares internacionales para la documentación, la difusión y la preservación de los metadatos y microdatos.

Data Documentation Initiative (DDI): Es un estándar basado en XML para la documentación de los microdatos. Proporciona un mecanismo simple para registrar y comunicar todas las características sobresalientes de los mismos.
Dublin Core Metadata Initiative (DCMI): Es un conjunto de elementos para la descripción de recursos digitales. Se utiliza para describir los recursos relacionados a los microdatos, como: cuestionarios, informes, manuales, scripts y programas de procesamiento de datos, etc.

2.1.1.2. Descripción de los elementos necesarios

Bases de datos (con o sin microdatos): Contiene las variables y registros de una operación estadística.

Diccionario de datos: Contiene información de la base de datos, archivo de donde viene la variable, el nombre del campo, la descripción del campo, el tipo y medida del dato y la longitud del campo.

Manual de diligenciamiento: Contiene la descripción de cada una de las preguntas del formulario y lo que debe contestarse en cada una de ellas.

Manual de validación y consistencia: Contiene información de la base de datos como el nombre de la tabla o archivo de donde viene la variable, el nombre del campo, la descripción del campo, el tipo de dato, la longitud del campo y adicionalmente, las reglas de validación y consistencia de la base de datos.

Material de referencia: Son los documentos de apoyo temático al proceso, pueden incluirse manuales, documentos de diseño, especificaciones de indicadores, guías, instructivos y presentaciones.

2.1.1.3. Elementos de documentación

Debe proveerse la información sobre las personas o entidades que intervinieron en la operación y la respectiva documentación. Se incluyen los elementos que describen las características del recurso que se enmarcan en 3 categorías generales: Metodología de la investigación, Descripción de las variables de la base de datos y Material de referencia.

2.2. INFORMACIÓN SOCIOECONÓMICA EN EL INSTITUTO HUMBOLDT

2.2.1. Técnicas utilizadas

La recopilación de información socioeconómica generada en el Instituto Humboldt se realizó con el fin de identificar qué tipo de información podría ser custodiada por la Infraestructura Institucional de Datos e Información (I2D) y se inició con el programa de Ciencias Sociales y Saberes de la Biodiversidad (CSSB), debido a que es él área que mayor cantidad de información de este tipo genera.

Al realizar el acercamiento con los investigadores del CSSB y al hacer la revisión de la información entregada y algunos documentos adicionales, se realizó un listado inicial con los instrumentos o técnicas utilizadas para la recopilación de información primaria que se pudieron identificar, los cuales son: Encuestas, entrevistas, talleres, recorridos, calendarios de producción, mapas parlantes (cartografía social), fichas técnicas de caracterización, análisis etnográfico, línea de tiempo, historia de vida, guías de observación etnográfica, diálogos semiestructurados, diagrama de finca y observación participante.

Posteriormente, el listado fue corroborado con ayuda de los investigadores y se estableció que algunos de los instrumentos identificados se complementan unos con otros. En el caso del desarrollo de un taller se pueden generar diagramas de finca, líneas de tiempo, mapas parlantes y calendarios de producción; la entrevista puede emplearse para hacer historias de vida y diálogos semiestructurados; un ejercicio de observación participante al igual que un análisis etnográfico pueden realizarse utilizando guías de observación; y, en un recorrido se pueden emplear guías de observación y fichas técnicas.

Con base en la depuración mencionada anteriormente, se determinó que las técnicas utilizadas en el desarrollo de las diferentes investigaciones son las siguientes:

1. Encuesta
2. Entrevista
3. Taller
4. Recorrido
5. Fichas técnicas (de observación, de caracterización)
6. Guías de observación

Los resultados obtenidos en cada una de las técnicas, son el insumo para realizar diversos análisis y para generar los resultados de los diversos estudios que son publicados, debido a esto, dichos instrumentos normalmente son sistematizados con el fin de facilitar los análisis, y cada uno se caracteriza porque tiene una dinámica particular para su desarrollo. Por ejemplo, en la entrevista y la encuesta la interacción es personal, ya que al aplicar esta técnica se debe establecer una comunicación individual entre el entrevistado y el entrevistador; en el taller la interacción es grupal, mientras que en los recorridos, las fichas técnicas y las guías de observación la información resultante corresponde a la interpretación que haga quien emplea esta técnica. Adicionalmente, cada técnica tiene un instrumento de recolección de información que es empleado que es sistematizado.

2.2.2. Acceso

Las mayores restricciones en cuanto al tipo de acceso, corresponden a las técnicas que emplean la comunicación individual, esto debido a que es muy común que el investigador establezca una relación de confianza con el entrevistado, en donde la sistematización de la información obtenida corresponde a las respuestas del entrevistador, es decir que lo registrado mediante estas técnicas son literalmente las respuestas de las personas. De otra parte, en muchos casos durante la aplicación de estas técnicas se manifiesta que el uso de la información recolectada será exclusivamente para el desarrollo de la investigación que se adelanta, es por esto que por defecto se restringe el uso de esta información tanto de manera interna como externamente y se deja abierta la posibilidad a que el investigador pueda cambiar el tipo de restricción siendo la persona responsable por el cambio de la restricción, sin embargo el acceso a los datos debe regirse según la Política Institucional de Datos, Información y Conocimiento una vez esta sea consolidada.

2.2.2.1. Tipos de acceso

Libres a nivel externo e interno: La I2D podrá compartir la información a entidades externas e investigadores del IAvH sin ninguna restricción.

Libres a nivel interno: La I2D podrá compartir la información únicamente a investigadores del IAvH sin ninguna restricción. Estos conjuntos de datos serán compartidos únicamente de manera interna, exclusivamente para uso institucional.

Libres a nivel interno con aviso previo: La I2D podrá compartir la información únicamente a investigadores del IAvH con previo aviso al investigador principal. Estos conjuntos de datos serán compartidos únicamente de manera interna, exclusivamente para uso institucional, previa aprobación del responsable.

Restricción temporal: La I2D se compromete a no compartir la información por un tiempo definido. El investigador principal debe determinar la fecha de liberación para que la I2D pueda compartir la información. Después de dicha fecha el tipo de acceso debe ser redefinido teniendo en cuenta los otros tipos de acceso.

3. PROPUESTA DEL PERFIL

Debido a la diversidad de la información en cuanto a contenido y formatos se dificulta enormemente la posibilidad de estructurar y estandarizar de manera similar a la realizada en otros componentes de la I2D (componentes biológico y geográfico) los datos derivados de cada investigación. Sin embargo la información recopilada puede ser incorporada de manera estandarizada con una apropiada descripción de cada estudio o investigación (metadatos), siguiendo el modelo implementado en el DANE. A pesar de implicar un tiempo extra de incorporación, la información podrá hacerse más visible y por lo tanto ser descubierta y utilizada por los investigadores que lo requieran.

Basado en los lineamientos para la documentación de metadatos a partir de los estándares DDI y Dublin Core, se propone la siguiente documentación para facilitar el descubrimiento de los datos e información socioeconómica (Tabla 2), se anexa la plantilla de documentación en formato Excel y el Diccionario de Datos en el mismo archivo.

Tabla 2. Secciones y campos de la plantilla de documentación de metadatos socioeconómicos.

DESCRIPCIÓN DEL DOCUMENTO	Los estudios o investigaciones no siempre son documentados por el mismo organismo o persona que produjo los datos e información. Es importante proveer en el metadato la información sobre las personas y/o entidades que intervinieron en él y durante el proceso de su documentación.
Responsables de la documentación	Nombre, afiliación y rol de los individuos y organizaciones involucradas en la documentación de la operación (no son necesariamente los productores).
Nombres y apellidos*	Nombres y apellidos completos. 🕮 **Nombre**

Afiliación*	Entidad a la cual se encuentra adscrita al momento de realizar la documentación. 🕮 **Institución**
Rol*	🕮 **Rol**
Fecha de documentación*	Fecha de la documentación, debe ser actualizada en cada revisión y modificación. Se recomienda usar el formato AAAA-MM-DD. 🕮 **Fecha**
Versión*	Permite hacer un seguimiento a la versión del documento, lo que facilita a los usuarios determinar si cuentan con la versión más reciente. Debe incluir los cambios realizados. Ejemplo: • **Versión 2 (Febrero de 2011). Esta versión reemplaza la versión 1 (enero de 2011) ya que incluye la versión final revisada por el supervisor del contrato.**
Identificador único*	El identificador para el metadato. El campo debe ser el identificador interno suministrado por la I2D al momento de la incorporación Ejemplo: • **I2D-OT_2014_IN001**
DESCRIPCIÓN DEL ESTUDIO	Esta consta de la información general del estudio o investigación; incluye la información sobre cómo debe ser citada, quién recolectó, compiló y distribuyó los datos, el contenido de los datos, los métodos de recolección, procesamiento, etc.
Identificación	Los elementos de este grupo se refieren a los estudios o investigaciones.
Título*	Escribir el nombre oficial completo, evitando el uso de abreviaciones.
Subtítulo	Título secundario si existe.
Tipo de estudio*	Según la temática y el método de recolección se selecciona el tipo de estudio que se presenta en el vocabulario controlado. 🕮 **Tipo de estudio**
Descripción general	Los elementos en este grupo proporcionan un resumen general sobre el estudio o investigación.
Resumen*	Idea general pero precisa de los diversos aspectos que componen el estudio. El objetivo es realizar un planteamiento claro y ordenado del tema de estudio, de su importancia e implicaciones.
Objetivos*	• **Objetivo general:** Es el propósito de la investigación. El objetivo general no es otra cosa que el título del proyecto y la pregunta de investigación planteados en términos de acción. • **Objetivos específicos:** Son expresiones cualitativas que deben conducir a los resultados.
Marco de referencia	El marco de referencia está compuesto por el marco teórico, el marco conceptual, el marco legal y aquellos temas que se considere necesario incluir, con la finalidad de contextualizar la investigación de la mejor manera posible.

Unidades muestreal (DANE: Unidad Estadística)	Las unidades estadísticas son el objeto del estudio. Estas unidades se dividen en unidades de observación, análisis y muestreo. 1. Unidad de observación: Es aquella unidad de la que se reciben los datos y se compilan estadísticas. Ejemplo: • Para un censo agropecuario la unidad de observación es: **predio**. 2. Unidad de análisis: Corresponde a la entidad objeto de estudio en una medición y sobre la que se presentan las conclusiones. Ejemplo: • Para una entrevista en una comunidad: la unidad de análisis es: **las personas que han sido enmarcados en la encuesta.** 3. Unidad muestral: Describe y define el elemento o conjunto de elementos a seleccionar en la muestra. Sólo aplica para operaciones por muestreo. Ejemplo: • Para la una encuesta de Convivencia y Seguridad Ciudadana: **10 viviendas en seleccionadas y ubicadas en una manzana con límites definidos.**
Clase de los datos*	**• Pérdida de la biodiversidad** **• Usos de la biodiversidad** **• Social y político** 🕮 **Clase**
Cobertura temática	Permite la documentación de los tópicos que fueron cubiertos por la operación estadística.
Universo de estudio	Describir el conjunto de unidades o individuos a los que se refiere el estudio o que constituyen la colectividad de interés y que satisfacen una definición común.
Población objetivo	Delimita la parte del universo sobre la cual es posible generar estimaciones debido a la exclusión de algunos elementos que no cumplen con la definición del universo, o por consideraciones prácticas de quien realiza el estudio.
Contenido temático*	Nombrar las variables (campos o elementos) más importantes del estudio.
Cuestionarios	De presentarse debe describir la estructura y composición del cuestionario o entrevista. Adicionalmente, documentar los objetivos de cada uno de los módulos o capítulos que componen el cuestionario, así como los objetivos de las preguntas.
Palabras clave	Relacionar o listar los conceptos fundamentales que permiten entender la investigación.
Cobertura geográfica	
País*	Especificar el país en que fue ejecutado el estudio. 🕮 **País**
Regiones administrativas	Definir las áreas geográficas del país que se cubren. 🕮 **Departamento** 🕮 **Municipio** 🕮 **Centro poblado**
Productores y auspiciadores	Contiene elementos que pueden ser utilizados para reconocer la labor de personas y organizaciones responsables del diseño, implementación y financiamiento del estudio.
Investigador principal*	Ingresar los datos del programa y persona responsable de la investigación. 🕮 **Nombre**
Colaboradores	Ingresar los datos de las personas u organizaciones involucradas en las distintas etapas: diseño,

	levantamiento de campo, procesamiento y análisis. Se recomienda seguir el formato: Nombres y apellidos, afiliación, rol. 🕮 **Nombre**
Financiamiento	Ingresar los datos de las fuentes de financiamiento (recursos fiscales y/u otras fuentes). Si existe, especificar el número de donación o contrato.
Reconocimientos	Ingresar los datos de las personas y/o instituciones que contribuyeron: asistencia técnica, estudios de referencia, etc.
Recolección de datos	Este grupo de elementos tiene la intención de proveer de un breve resumen de las actividades de recolección de los datos.
Periodo de recolección	Ingresar las fechas o rango de fechas en que se realiza la recolección de los datos e indicar el ciclo (mensual, trimestral, anual, etc.). En formato AAAA-MM-DD/AAAA-MM-DD, AAAA-MM o AAAA. 🕮 **Fecha**
Notas sobre la recolección de datos	Describir aspectos importantes que se hayan presentado durante la recolección de datos. Incluir factores tales como cooperación de los encuestados, duración de las entrevistas, número de visitas y otros hechos y ocurrencias.
Supervisión y control del operativo de recolección	Presentar los controles que se aplican en terreno al personal a cargo o coordinador de campo, tales como: medición de rendimientos, recorridos, cobertura de las áreas geográficas y de unidades asignadas, equipos y dispositivos; así como controles para reducir el sesgo.
Recolector de datos	Proporciona un listado de las personas y las organizaciones responsables de recolectar los datos. Esto se refiere a quien recolecta los datos, no a las que producen la documentación. 🕮 **Nombre**
Procesamiento de datos	En este apartado se ingresa un resumen general de los procedimientos utilizados para identificar y corregir errores en los datos.
Depuración de datos	El ambiente, las herramientas y la arquitectura propuestas o utilizadas para el procesamiento de datos. Las etapas que se llevan a cabo en el procesamiento de datos (según el caso, transmisión de datos, controles de calidad, consolidación de archivos, validación y verificación de la consistencia interna de los datos, imputación, generación de resultados, seguridad y almacenamiento de la información). Se debe también proporcionar el conjunto de reglas o especificaciones de validación, consistencia e imputación.
Acceso a las bases de datos	Resumen de las condiciones bajo las cuales se puede acceder a los datos de un estudio.
Institución propietaria	Ingresar los datos de la institución propietaria de la base de datos y los resultados.
Condiciones de acceso	• **Libres a nivel externo e interno** • **Libres a nivel interno** • **Libres a nivel interno con aviso previo** 🕮 **Condición de acceso**
Requerimientos de citación	Especificar el texto para acreditar la fuente de datos conforme a las normas de citación, siga el formato: Entidad (año). Titulo. Tipo. Nombre 1 (Rol 1). Nombre 2 (Rol 2) Versión #. URL • **Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (2017). Actores clave en los bosques montanos del Oriente Antioqueño, Cañón del río Melcocho, en el municipio de El Carmen de Viboral. Entrevista. Barriga, J. (Contacto del recurso), Acosta, A. (Creador del recurso), Quintana, A. (Proveedor de metadatos). Versión 3.0. http://i2d.humboldt.org.co/sec**

Notas legales y derechos de autor	
Notas legales	Proveer información relativa a la responsabilidad de los usuarios de la documentación y las bases de datos.
Derechos de autor	Proveer la declaración de derechos de autor de la documentación y las bases de datos. Se recomienda hasta donde sea posible reutilizar licencias estándar como las creative commons, *copyleft* o *copyright*.
Contactos	
Personas a contactar	Ingresar el nombre de las personas responsables de la documentación y la base de datos. Especificar correo electrónico. 🕮 **Nombre**
BASE DE DATOS	Esta sección contiene todos los archivos de datos asociados, además de una descripción detallada de cada variable. Los elementos en esta sección son obtenidos de las especificaciones DDI.
Descripción del archivo	
Contenido	Describir la base de datos incluyendo cobertura temática, características especiales de su contenido y principales variables.
Productor	Ingresar el nombre de la institución y persona responsable de la creación de la base de datos (el productor de la base de datos no es necesariamente la quien que recolectó o procesó los datos). 🕮 **Nombre**
Versión	Describir cada versión de la base de datos. Incluir fecha y detalles de ajustes realizados a la base de datos.
Comentarios del archivo	Ingresar comentarios adicionales acerca del archivo de datos, como información general acerca de la causa por la que se perdieron datos. Ejemplo: fusión de archivos, error de programación, etc.
Campos o elementos	
Listado de campos	De presentarse, proveer el nombre de las variables y su etiqueta.
Descripción de los campos	Documentar el grupo de personas u otros elementos objetivo de la variable y a los que se refieren los resultados analíticos.

3.1. HERRAMIENTA DE PUBLICACIÓN

Adaptado de *Norwegian Centre for Research* Data http://www.nesstar.com)

Se propone como herramienta Nesstar, desarrollada por el *Norwegian Centre for Research*, Nesstar es un sistema de software para publicación de datos y análisis en línea compatible e interoperable con el Archivo Nacional de datos del DANE, entidad de referencia para la gestión de datos e información socioeconómica. El software consiste en herramientas que permiten a los proveedores de datos hacer visibles sus datos e

información. Nesstar maneja datos de encuestas y tablas multidimensionales, así como recursos de texto. Los usuarios pueden buscar, examinar y analizar los datos en línea.

3.1.1. Nesstar Publisher

Nesstar Publisher es un programa de gestión de datos. Puede utilizarse como una herramienta independiente para la preparación de metadatos, pero también puede utilizarse para montar el conjunto de datos mejorado en un servidor Nesstar.

Publisher está disponible sólo en Windows. La única manera de ejecutarlo en otra plataforma como Mac / Linux / etc sería a través de Windows (a través de una máquina virtual). Es gratis y puede ser descargado y utilizado por cualquier persona sin licencia. La versión gratis no cuenta con soporte.

3.1.1.1. Características principales

- Fácil edición y creación de conjuntos de datos documentados de DDI sin necesidad de conocer XML.
- Capacidad de importar y exportar solo archivos DDI.
- La capacidad de crear plantillas que permitan a su organización estandarizar su uso del DDI.
- Agregar texto predeterminado y vocabularios controlados a las plantillas
- La capacidad de crear grupos de variables
- La inclusión de las opciones de frecuencia y estadísticas de resumen para cada variable.
- La capacidad de importar y exportar datos a los formatos estadísticos más comunes.
- La capacidad de insertar enlaces a sitios web o documentos relevantes.
- Funcionalidad para administrar recursos publicados en un servidor Nesstar.
- La capacidad de vincular conjuntos de datos jerárquicamente relacionados.

3.1.1.2. Requerimientos

Publisher es una aplicación independiente de escritorio. Requiere Windows XP, 2000, Vista o Windows 7 y conexión a internet si los datos se van a publicar directamente en el servidor Nesstar.

3.1.2. Nesstar Server

Nesstar Server ofrece la capacidad de publicar información que los usuarios pueden buscar, examinar, analizar y descargar. El software está disponible para sistemas operativos Windows y Linux.

3.1.2.1. Características principales

Publicación de estadísticas en el servidor: Nesstar está construido como una extensión para un servidor web normal. Además de proporcionar todas las facilidades habituales para la publicación de contenido web, este servidor ofrece la posibilidad de publicar información estadística que los usuarios pueden buscar, examinar, analizar y descargar. Esto se hace mediante el uso de un navegador web estándar que utiliza Nesstar WebView, la interfaz web de Nesstar Server.

Combinación de datos y metadatos: Nesstar proporciona una forma de publicar, ubicar, acceder y proteger contra el acceso no autorizado. A través de este proceso, el conocimiento se hace más accesible. Esto incluye la combinación de metadatos, datos e información derivada (informes, documentos, análisis secundarios, etc.).

Una web de datos semánticos: Nesstar representa un sistema de arquitectura de software que facilita la creación, localización, acceso y operación remota en metadatos y datos correspondientes. Al mismo tiempo, lo hace manteniendo un alto nivel de compatibilidad con la WWW. Este es un intento de crear una Web de datos; una extensión de la WWW orientada a datos y basada en estándares de descripción de datos ampliamente adoptados.

Sistema basado en estándares: Las demandas de sistemas reconocidos como la Iniciativa de Documentación de Datos (DDI) hacen que el proceso de usar Nesstar sea una operación completamente integrada y estandarizada.

Autenticación y control de acceso: El acceso a los recursos puede ser gratuito y universal, o puede limitarse a los usuarios autorizados de acuerdo con los deseos del editor de datos. Estos materiales están listos para ser usados y explorados usando Nesstar WebView.

3.1.2.2. Requerimientos

Para Windows:

- 2 GB de RAM (4 GB es preferible).
- 2Ghz CPU o más rápido.
- Disco duro de 10 GB como mínimo.

Para LInux:

- Linux (ejecutándose en i386 o procesadores compatibles)
- OpenJDK 1.7
- MariaDB
- Apache Ant> = 1.7

5. REFERENCIAS

ANDA, Inicio. (n.d.). Retrieved August 29, 2017, from http://formularios.dane.gov.co

Murdoch, W., Polasky, S., Wilson, K. A., Possingham, H. P., Kareiva, P., & Shaw, R. (2007). Maximizing return on investment in conservation. *Biological Conservation*, *139*(3-4), 375–388.

Polasky, S. (2008). Why conservation planning needs socioeconomic data. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(18), 6505–6506.

Wetzel, F. T., Saarenmaa, H., Regan, E., Martin, C. S., Mergen, P., Smirnova, L., … Häuser, C. L. (2015). The roles and contributions of Biodiversity Observation Networks (BONs) in better tracking progress to 2020 biodiversity targets: a European case study. *Biodiversity* , *16*(2-3), 137–149.