

GUÍA DEL ADMINISTRADOR DE INFORMACIÓN SOBRE BIODIVERSIDAD

Ángela M. Suárez-Mayorga
Editora

Sistema de Información sobre Biodiversidad de Colombia - SiB
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt
Entidad Coordinadora



INSTITUTO DE INVESTIGACIÓN DE RECURSOS BIOLÓGICOS
ALEXANDER VON HUMBOLDT

(CC) Instituto de Investigación de Recursos Biológicos Alexander von Humboldt
2007

Se permite reproducir y comunicar esta obra siempre y cuando se cite la fuente de manera correcta y no se utilice para fines comerciales sin la previa autorización del titular. Toda obra derivada deberá utilizar el mismo tipo de licencia Creative Commons que respalda la presente obra.

Algunos derechos reservados

<http://creativecommons.org/licenses/bync-sa/2.5/co>

EDICIÓN:

Claudia María Villa G.

FOTOGRAFÍAS:

Francisco Nieto M.

Banco de Imágenes Ambientales, Instituto Alexander von Humboldt

Contribución IAvH No. 409

DIAGRAMACIÓN E IMPRESIÓN: ARFO Editores e Impresores Ltda.

Impreso en Bogotá, Colombia

Diciembre de 2007

Para citar la obra completa, cítese como: Suárez-Mayorga A.M. (ed.). 2007. Guía del administrador de información sobre biodiversidad. Sistema de Información sobre Biodiversidad de Colombia –SiB–, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D.C., Colombia, 74 p.

Para citar un capítulo, o un numeral específico, cítese como: Benavides Molineros J. 2007. Herramientas de evaluación de calidad, p. 35-38. En: Suárez-Mayorga A.M. Guía del administrador de información sobre biodiversidad. Sistema de Información sobre Biodiversidad de Colombia, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D.C., Colombia, 74 p.

ISBN: 978-958-8343-07-5

Palabras Clave: 1. Gestión de información. 2. Sistemas de información. 3. Registros biológicos. 4. Informática de la biodiversidad. 5. Colombia

El Sistema de Información sobre Biodiversidad de Colombia (SiB) es una iniciativa de carácter nacional, liderada por el Instituto Humboldt y conformada por múltiples entidades del país. Ha sido desarrollada para facilitar la gestión de datos e información que apoyen oportuna y eficientemente procesos de investigación, educación o toma de decisiones relacionadas con la diversidad biológica de Colombia. La **Guía del administrador de información sobre biodiversidad** hace parte de una serie de publicaciones elaboradas por el Equipo Coordinador del SiB para apoyar los procesos de creación de capacidad en torno a la iniciativa en Colombia y el mundo.

Comité Técnico Directivo del SiB (CTD)

Miembros: Instituto Amazónico de Investigaciones Científicas - SINCHI. Instituto de Ciencias Naturales, Universidad Nacional de Colombia. Instituto de Hidrología, Meteorología y Estudios Ambientales - IDEAM. Instituto de Investigaciones Ambientales del Pacífico - IIAAP. Instituto de Investigaciones Marinas y Costeras José Benito Vives de Andreis - INVEMAR. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt - IAvH

Invitado permanente: Ministerio de Ambiente, Vivienda y Desarrollo Territorial



Esta publicación ha sido realizada por el Instituto Humboldt en el marco del proyecto “Conservación y Uso Sostenible de la Biodiversidad en los Andes colombianos”, financiado por el Fondo Mundial Ambiental GEF, el Banco Mundial y la Embajada Real de los Países Bajos. La impresión de la obra fue financiada con recursos del Ministerio de Ambiente, Vivienda y Desarrollo Territorial.



FERNANDO GAST HARDERS
Director General

Autores

Nelson Andrés Ramón-Rodríguez

Ingeniero de sistemas, especialista en animación digital
Investigador junior – Desarrollador SiB 2003-2006
ponckk@yahoo.com

Ángela M. Suárez-Mayorga

Bióloga MSc
Investigadora senior – Administradora líder de contenidos SiB
amsuarez@humboldt.org.co

Diego Trujillo-Motta

Biólogo
Investigador junior – Administrador de AAT SiB
dmtrujillo@humboldt.org.co

Héctor Fabio Rivera Gutiérrez

Biólogo
Investigador junior – Administrador de metadatos SiB 2002-2005
hfrivera@gmail.com

Julia Benavides Molineros

Bióloga MSc
Investigadora senior – Productos y servicios de información SiB
jebenavides@humboldt.org.co

Talía Waldrón H.

Bióloga MSc
Investigadora junior – Administradora de tesauros SiB 2002-2005
tallia@rocketmail.com

Ximena Franco Villegas

Ingeniera industrial MSc
Investigadora principal SiB
xfranco@humboldt.org.co

Tabla de contenido

Presentación	7
Prólogo	9
I. Propósito y alcances	11
II. ¿Cómo hacer uso de esta guía?	13
III. Marco de referencia	15
i. El SiB como sistema distribuido	15
ii. Red de administradores de información sobre biodiversidad	16
IV. ¿Qué es un administrador de información?	19
i. Definición	19
ii. Perfil	19
V. ¿Cuál es la labor de un administrador en el marco del SiB?	21
i. Funciones	21
ii. Derechos	22
a. Administradores básicos	22
b. Administradores avanzados	23
iii. Deberes y responsabilidades	23
a. Administradores básicos	24
b. Administradores avanzados	24
VI. ¿Qué información se administra?	25
i. Datos vs. información	25
ii. Los objetos de interés en el marco del Sistema: entidades y atributos	27
VII. ¿Cómo se obtiene la información?	29
i. La captura de datos	29
ii. Calidad y consistencia	30
iii. Herramientas de evaluación de calidad	31
a. Estructuras y conjuntos de datos centrales	31

b.	Documentos de la estructura del Sistema	31
c.	Criterios de evaluación de calidad	33
■	<i>Criterios de evaluación de datos y conjuntos de datos</i>	33
■	<i>Criterios para la evaluación de información</i>	33
VIII.	¿Qué significa “datos administrados en el marco del SiB”?	35
i.	Estandarización: datos interoperables.	35
a.	Estándares	35
b.	Conjuntos de referencia	38
•	Archivos de autoridad taxonómica	39
•	Tesauro	40
•	Catálogo de métodos y atributos	41
c.	Procedimientos asociados	43
⇒	Consultas de AAT	43
⇒	Descargas de AAT	44
⇒	Verificación de los nombres	45
⇒	Descripción de un conjunto de datos mediante el tesauro	45
⇒	Búsquedas en el Sistema mediante el tesauro	46
⇒	Consultas al catálogo de métodos y atributos	47
⇒	Documentación de métodos y atributos	48
ii.	Documentación: datos multipropósito	50
a.	Catálogo de metadatos	50
b.	Procedimientos asociados	51
■	Documentación de metadatos	51
IX.	¿Cómo se almacena la información bajo los parámetros SiB?	55
i.	Registros biológicos	55
ii.	Estructuras recomendadas de base de datos	56
iii.	Seguridad y almacenamiento físico	60
iv.	Procedimientos asociados	61
X.	¿Cómo fluye la información?	63
i.	Definición de condiciones de acceso y uso	63
ii.	Acuerdos y convenios	64
iii.	Productos y servicios de información	66
XI.	Literatura citada	71

Presentación

Hace apenas cuatro años, cuando el Sistema de Información sobre Biodiversidad de Colombia –SiB– estaba comenzando a implementarse como una iniciativa nacional, Colombia empezó a hablar sobre estándares y procedimientos para describir, almacenar y compartir información biológica. De hecho, reconociendo la información como activo de alto valor estratégico en las organizaciones y en las comunidades, así como el rol de la biodiversidad como recurso y patrimonio de la Nación, el Equipo Coordinador del SiB en el Instituto Humboldt produjo en 2005 la primera serie de estándares nacionales para administrar información relacionada con biodiversidad. Este esfuerzo, que actualmente va por su segunda edición, representó un arduo trabajo de investigación, concertación y difusión, que fue posible gracias a la colaboración (o coautoría) de pares académicos de diferentes universidades e institutos de investigación, quienes creyeron en la iniciativa abierta y participativa planteada por el SiB entonces, bajo el liderazgo de Instituto Humboldt.

No obstante, la definición de reglas de juego claras para interactuar en torno a la información sobre biodiversidad no ha sido suficiente para que quienes la necesitan puedan aprovecharla de quienes la producen. Ha sido necesario diseñar, de nuevo de manera concertada, esquemas de capacitación, herramientas tecnológicas y procedimentales, y productos y servicios de información que hagan posible esa conexión necesaria entre generadores y usuarios. Pero sobre todo, ha sido necesario acompañar a estos involucrados a lo largo de todo el ciclo de gestión de la información, contribuyendo a inducir un cambio de mentalidad en unos y otros, que hoy día ha permitido que algunas decisiones que afectan la conservación de nuestra biodiversidad se tomen con argumentos reales, y que existan, a la fecha, cerca de un millón de registros de ocurrencias de organismos colombianos disponibles en Internet con público acceso y uso.

Para soportar este acompañamiento y, de alguna manera, garantizar la perdurabilidad de muchos de los conceptos que el SiB promueve, se diseñó la guía que aquí se presenta. Está enfocada explícitamente a los administradores de información –quienes la generan, la custodian y la ponen a disposición–, confiando en que sirva de apoyo a su incorporación en el esquema de gestión de información planteado por el SiB, pero también, en que se constituya en una herramienta de consulta y base para el desarrollo de sus procesos individuales.

Fernando Gast Harders
Director General

Prólogo

Durante décadas en Colombia se han estado adelantando procesos de investigación sobre diferentes aspectos relacionados con la biodiversidad marina y continental de nuestro país; como resultado de estos procesos se ha generado una cantidad aún no determinada de literatura científica, bases de datos, cartografía y material documental (por citar algunos), los cuales cubren una amplia gama de énfasis temáticos, así como una gran cantidad de grupos de organismos y regiones geográficas del país.

Si bien se intuye que la cantidad de información actualmente existente sobre la diversidad biológica de Colombia es enorme, mucha de esta información se encuentra dispersa, no es fácil de ubicar, está pobremente documentada y por lo general es almacenada en condiciones físicas y lógicas poco favorables. Esta situación hace que el potencial de aprovechamiento de los estudios relacionados con biodiversidad –y en particular de los datos y la información como algunos de sus resultados más tangibles– sea muy limitado, lo que a su vez reduce las posibilidades de apoyar adecuadamente (y por ende, de impactar positivamente) las actividades relacionadas con la conservación y el uso sostenible de la biodiversidad de nuestro país.

Reflexiones de este tipo influyeron la forma como se ha venido diseñando e implementando el Sistema de Información sobre Biodiversidad de Colombia (SiB), iniciativa liderada por el Instituto Humboldt en asocio con múltiples entidades del orden nacional, regional y local. Esta iniciativa busca promover y facilitar los procesos de construcción, uso y distribución del conocimiento sobre la biodiversidad de Colombia, principalmente mediante el establecimiento y puesta en marcha de estrategias para la gestión integral de los datos y la información sobre biodiversidad.

Resulta fundamental para el funcionamiento y la viabilidad del SiB, desarrollar mecanismos que vinculen a los diferentes actores involucrados en las actividades de investigación sobre la diversidad biológica de Colombia; de esta forma se espera que dichos actores participen activa y coordinadamente en la generación, el procesamiento, la disseminación o el intercambio de datos e información sobre biodiversidad, todo esto en el marco de procesos de investigación con visión estratégica, de mediano-largo plazo y respondiendo –en la medida de lo posible– a las necesidades del país en esta materia.

En este contexto ha surgido la idea de fortalecer la capacidad nacional para la gestión de datos e información sobre biodiversidad, mediante estrategias como la formación de administradores de información. Un administrador es, en este contexto, una persona, un grupo o una entidad que produce o custodia datos o información sobre la biodiversidad del país siguiendo las pautas definidas y acordadas colectivamente en el marco del SiB; en este sentido, un administrador mantiene la autonomía en el manejo de sus

propios datos, aunque se identifica con –y por lo tanto asume– la visión de la información como un recurso estratégico para conservar y usar sosteniblemente la biodiversidad, y que como tal debe ser cuidado, compartido y aprovechado.

La presente guía responde a este enfoque, sirviendo no solamente para la formación de administradores de datos e información, sino además para promover la participación de múltiples entidades y personas en los procesos de construcción y utilización del conocimiento sobre la biodiversidad de nuestro país. Usando un lenguaje sencillo y amigable, esta guía brinda los elementos técnicos y conceptuales para entender y asumir el rol de administrador en el marco del SiB. La primera versión (2004) ha sido enriquecida y complementada con todos los comentarios y aportes de quienes la han usado como herramienta de apoyo en su proceso de vinculación al SiB. La actual versión recoge estos aportes y se extiende en los nuevos desarrollos del Sistema, con un mayor énfasis en los productos y servicios de información actualmente disponibles. Este documento es el resultado del trabajo adelantado por el equipo coordinador del SiB en su objetivo por consolidarse como una alianza nacional que facilite la gestión de información en pro de la conservación y el uso sostenible de la diversidad biológica de Colombia.

Juan Carlos Bello Silva
Oficial de Nodos
Global Biodiversity Information Facility - GBIF
Investigador asociado SiB

I. Propósito y alcances

Esta guía fue construida para apoyar la labor de los *administradores de datos e información* en el marco del Sistema de Información sobre Biodiversidad de Colombia (SiB). Por lo tanto, se espera que sirva de base conceptual para entender el esquema de administración de contenidos del Sistema y que proporcione respuestas a las preguntas que puedan surgir durante los procesos de captura, documentación y puesta a disposición de datos e información que un administrador lleve a cabo.

Específicamente, la guía tiene por objeto:

- ~ describir los elementos conceptuales básicos a tener en cuenta cuando se gestiona información a través del SiB,
- ~ instruir a los administradores sobre procesos y procedimientos necesarios en la incorporación de sus datos e información al Sistema y, finalmente,
- ~ sustentar las capacitaciones que los administradores reciben por parte del Equipo Coordinador del SiB (EC-SiB) cuando se vinculan como tales.

Como complemento a esta guía se encuentran los estándares y guías de usuario desarrollados por el componente de administración de contenidos del EC-SiB, así como la *Política institucional para la gestión de información sobre biodiversidad* (Franco Villegas 2003). Se espera que esta guía facilite también la lectura de estos documentos y contribuya a, aclarar sus contenidos y divulgarlos de una manera didáctica.

II. ¿Cómo hacer uso de esta guía?

Cada uno de los capítulos de la guía está identificado con un numeral romano en mayúsculas. El título de los capítulos fue diseñado para que el usuario haga la búsqueda partiendo de las preguntas que se pueda formular; por lo tanto es necesario que, una vez planteada la pregunta, el usuario decida el aspecto que desea detallar, a través de una consulta en la tabla de contenido. Los contenidos detallados se encuentran representados por subtítulos de segundo, tercer y cuarto nivel.

El nombre de los autores de las secciones y subsecciones dentro de cada capítulo se encuentra en el extremo derecho de la línea de título. En los capítulos, secciones y subsecciones donde no aparece un nombre de autor, debe entenderse que el autor es la editora del documento general.

En cuanto a los contenidos, existen unas convenciones generales de escritura que hacen referencia a conceptos especiales:

- ~ Los conceptos clave que cuentan con una definición inmediata se representan mediante términos en **negrita** cuando se definen por primera vez, y en *itálica* cuando se mencionan posteriormente, siempre que sea necesario resaltarlos.
- ~ Los conceptos relevantes que no cuentan con una definición inmediata se resaltan en *itálica*.
- ~ Los conjuntos de datos de referencia son equivalentes a conceptos clave en el texto, por lo que se representan en **negrita** cuando se definen por primera vez y en *itálica* cuando se hace mención de ellos.
- ~ Cuando los conjuntos de datos son descritos en detalle, se presentan subrayados y resaltados con una viñeta circular.
- ~ Los corolarios que no son definiciones se resaltan con un marco.
- ~ Los procedimientos se anotan mediante viñetas rectas con sangría amplia.
- ~ Los requisitos, condiciones y criterios de personas, organizaciones o entidades se anotan con viñetas cuadradas.

- ~ Los compromisos de personas e instituciones con el sistema o del sistema con sus usuarios se resaltan con el símbolo ✓.
- ~ En el apartado de productos y servicios (capítulo XI), se utiliza un asterisco para señalar los pasos comunes al desarrollo de cualquier producto de información.

Por último, en algunas secciones del texto se hace referencia a contenidos presentes en otra sección o título del mismo documento, con el fin de que el usuario los recuerde o los detalle. Estas referencias se representan en *itálica*, precedidas del símbolo “→” y entre paréntesis, y están identificadas claramente en la tabla de contenido. Es ideal entonces que el usuario consulte la tabla de contenido para ubicarlas rápidamente.

III. Marco de referencia

i. El SiB como sistema distribuido

Ángela M. Suárez-Mayorga y Ximena Franco-Villegas

El Sistema de Información sobre Biodiversidad de Colombia (SiB) fue creado como respuesta a los compromisos adquiridos por el país con la ratificación del Convenio sobre Diversidad Biológica (CDB) y en consecuencia con el Código de Recursos Naturales (Decreto 2811 de 1974) y con la Ley 99 de 1993, los cuales mencionan y determinan la creación del Sistema de Información Ambiental para Colombia (SIAC), donde la biodiversidad es uno de los tópicos principales. El desarrollo y la coordinación del SiB fueron encomendados al Instituto de Investigación de Recursos Biológicos Alexander von Humboldt desde el momento mismo de su nacimiento –decreto 1603 de 1994–. Sin embargo, su concepción partió también de la necesidad de rescatar y aprovechar el trabajo previo y la experiencia de otras personas y entidades en cuanto a gestión de información sobre biodiversidad, como los demás institutos de investigación del Sistema Nacional Ambiental (SINA) y la comunidad académica y científica.

En consecuencia, durante 2002 se comenzó a trabajar la propuesta de un SiB nacional, convocando a aquellas personas y entidades cuya tradición y conocimiento podían dar soporte al proceso. Más tarde se formuló la propuesta de un **Comité Técnico Directivo (CTD-SiB)**, conformado por los directores de las entidades directamente involucradas y se desarrolló una versión preliminar de un documento formal sobre el sistema de información sobre biodiversidad en la escala nacional.

De acuerdo con Bello (2002 citado en Franco Villegas 2003), un sistema de información sobre biodiversidad pertinente y útil para un país como Colombia debería contar con las siguientes características:

- Permitir la gestión de información de manera abierta y distribuida, considerando la participación de múltiples personas y organizaciones espacialmente separadas.
- Basarse en estándares y protocolos, en la medida de lo posible validados en el contexto internacional.
- Promover el uso de tecnologías que faciliten el intercambio y la integración de información.
- Ser flexible en el desarrollo de contenidos, particularmente en relación con la definición de necesidades y prioridades de información.
- Atender demandas específicas en términos de productos o servicios de información, considerando una definición previa y clara de los públicos objetivo.

Para hacer frente a los requerimientos arriba expresados, fue claro que la arquitectura del sistema propuesto debía ser **distribuida**, es decir, una red de redes donde en lugar de compilar información en una sola entidad, cada conjunto de datos integrado debía ser administrado por su generador o custodio (diferentes personas y entidades), y en la que los procesos y responsabilidades de gestión de información se distribuyeran entre sus miembros, teniendo como pilar fundamental la creación de capacidad individual, institucional y sistémica. En este esquema, sin embargo, se hizo necesaria la designación de un **Equipo Coordinador (EC-SiB)**, que hace las veces de una secretaría técnica, acoge e implementa las recomendaciones del *Comité Técnico Directivo (CTD)* y, por mandato de ley, se encuentra alojado en el Instituto Humboldt. Aunque ha sido necesario implementar varios desarrollos conceptuales y tecnológicos para que efectivamente se provea información en el esquema distribuido, en el mundo actual resulta la arquitectura más costo-eficiente, pues garantiza también la distribución de costos de participación. Prueba de lo anterior es que otras redes regionales y globales de gestión de información (no únicamente de información biológica) están implementando el mismo modelo.

Según Bello *et al.* (2003), un sistema distribuido busca “la consolidación de procesos de gestión abierta, dinámica y multidireccional, que favorezcan el intercambio y la comunicación directa entre los proveedores y los usuarios de los datos y la información sobre biodiversidad”, siempre con base en la creación de capacidad y confianza. En efecto, se ha visto que un esquema de este tipo facilita la integración de aquellas personas o entidades cuyo trabajo tiene especial relevancia en el marco del Sistema: permite armonizar los sistemas y redes ya existentes, utilizar tecnologías de punta e intercambiar información en un marco de calidad, seguridad y confianza, a través del uso de estándares y protocolos validados por la comunidad.

Es entonces bajo el esquema de “red de redes” o “sistema de sistemas” que se formula la propuesta SiB: partiendo de la creación de capacidad, se busca a través de herramientas como la presente guía, los conjuntos de datos de referencia, o los estándares para describir, organizar y compartir información, identificar, formar y fortalecer redes de administradores de información, que sean lo suficientemente sólidas como para gestionar su información con calidad y pertinencia, tanto para sus propósitos específicos como para los propósitos del Sistema.

ii. Red de administradores de información sobre biodiversidad

Coaliciones, alianzas, asociaciones, son términos relacionados con las redes. Sin embargo, la particularidad del término **redes** es que hace referencia al *tema de la información* y de esta forma se establece una diferencia con otros tipos de agrupaciones que sencillamente se generen en torno a temas o actividades específicas, sin que ello implique la estructuración de relaciones a partir de la información.

La conformación de las redes se da en torno a temas e *intereses comunes* o a la integración de perspectivas; sin embargo, no se puede dejar de lado que cada integrante de una red representa intereses diferentes a los comunes, pero que también son llevados a la red, generando de esta forma, al mismo tiempo, un panorama de *múltiples intereses*.

Es necesario reconocer que la participación de los actores en redes se da a partir de una motivación por el *valor agregado* que la pertenencia a la red represente. No existirán dinamismo y flujos de información en

el ámbito de una red si ésta no significa una mejora en algún aspecto (desempeño, conocimientos) de una organización. Una red puede conformar, en unión con otras, redes más amplias, generando nodos o subsistemas, con los cuales es posible modular¹ esos diferentes niveles de agregación o jerarquización de la red, para obtener diversos resultados.

De acuerdo con la *Estrategia de gestión interinstitucional para la implementación del Sistema de Información sobre Biodiversidad* (Franco Villegas y Bello 2004), los *administradores* (→ ¿Qué es un administrador de información?) son también usuarios del sistema, es decir, al generar y administrar información, utilizan la infraestructura, los productos, los métodos, los procedimientos, los estándares, las normas y en general, todas las herramientas desarrolladas y administradas por el EC-SiB para la gestión de información en el ambiente del SiB. En ese contexto, se define como **red de administradores** el conjunto de personas, grupos o entidades agrupadas en torno a la gestión de información sobre biodiversidad. En el marco del Sistema, hay dos tipos de redes de administradores: **redes geográficas regionales**, que agrupan administradores distribuidos físicamente en un ámbito geográfico restringido y definido (como por ejemplo la red de administradores del Eje Cafetero), y **redes temáticas**, que agrupan administradores alrededor de un tópico común, como es el caso de la Red Nacional de Observadores de Aves (RNOA).

Un **replicador de capacidad**, por su parte, es una entidad dentro de la *red regional o temática de administradores* que, aprovechando su capacidad de convocatoria y liderazgo regional o temático, se posiciona para difundir los conceptos y herramientas para la implementación del SiB, a la vez que incorpora el esquema de gestión de información del SiB dentro de sus propios procesos con el acompañamiento directo del EC-SiB. En términos de sistemas, estas son *redes anidadas* o *subsistemas* haciendo parte del SiB y replicando sus contenidos y herramientas a escala.

Con las redes geográficas y temáticas de administradores se busca multiplicar la capacidad y posibilidades del EC-SiB para replicar de la forma más amplia y eficiente posible la implementación del SiB a nivel nacional; se procura construir subsistemas de información sobre biodiversidad a fin de cubrir las necesidades de información en el país a escalas más pequeñas de lo que el Equipo Coordinador del SiB individualmente podría lograr.

Son *compromisos* de los replicadores de capacidad los siguientes:

- ✓ Implementar el esquema SiB para la administración de sus propios datos e información sobre biodiversidad.
- ✓ Apoyar a los miembros de la *red de administradores* en el desarrollo de su capacidad institucional para la gestión de datos e información sobre biodiversidad, mediante capacitación y acompañamiento a las entidades en la implementación de procesos institucionales.
- ✓ Llevar a cabo la coordinación interinstitucional en términos de información sobre biodiversidad entre las entidades *administradoras*.
- ✓ Liderar la gestión de fondos para la continuidad de la *red de administradores* al nivel que le corresponde.
- ✓ Participar en actividades del EC-SiB que agrupen a las entidades *replicadoras de capacidad* del país.

¹ **Modular**: modificar los factores que intervienen en un proceso para obtener distintos resultados. Tomado del Diccionario de la Real Academia de la Lengua Española, RAE, 2001.

Son *compromisos* de los administradores, al formar parte de una red, los siguientes:

- ✓ Participar en las actividades de capacitación y acompañamiento que lleve a cabo la *entidad replicadora de capacidad* correspondiente.
- ✓ Apoyar a las demás entidades que conforman la *red de administradores* en la implementación de procesos regionales o temáticos, según el esquema bajo el cual se agrupe la red.

IV. ¿Qué es un administrador de información?

i. Definición

De acuerdo con el *Esquema general de membresías del Sistema de Información sobre Biodiversidad* (Franco Villegas 2004) y la *Estrategia de gestión interinstitucional para la implementación del Sistema de Información sobre Biodiversidad* (Franco Villegas y Bello 2004), un **administrador de información** en el marco del SiB es una persona o entidad que genera o gestiona información sobre biodiversidad, utilizando la infraestructura, los productos, los métodos, los procedimientos, los estándares, las normas y, en general, las herramientas desarrolladas o sugeridas por el Equipo Coordinador del SiB (EC-SiB), con el fin último de construir conocimiento sobre biodiversidad.

El administrador es entonces un *involucrado* en la gestión de datos e información y un socio del SiB; como tal entra a formar parte de la *red de administradores* de datos e información en el Sistema (Franco Villegas 2003). La vinculación como *administrador* se da por voluntad de la entidad o persona administradora y por la invitación directa del EC-SiB. La calidad de *administrador* puede ser temporal, bien porque no se mantengan ni actualicen los conjuntos de datos en el tiempo o bien porque se establecen contactos esporádicos con el Sistema para documentar, estructurar o poner a disposición pública eventualmente conjuntos de datos.

ii. Perfil

El desempeño como administrador de datos e información en el marco del SiB requiere de personas con calidades humanas y profesionales sobresalientes (siendo éste por definición un concepto subjetivo), ya que de la responsabilidad con que se asuma esta función dependerán la calidad y pertinencia de la información sobre biodiversidad que se incorpora al Sistema. Dado que el SiB es el único sistema de información sobre biodiversidad formalmente establecido en el país, quienes se desempeñen como administradores de información (sin importar la temporalidad de la administración) deben ser conscientes de que la información a su cargo que sea gestionada a través del mismo debe ser de la mejor calidad² (→ *Calidad y consistencia*). Así mismo, debe responder a las necesidades de información del país y de sus múltiples usuarios y audiencias.

² Nótese que **calidad** en el contexto SiB no implica necesariamente veracidad en los contenidos sino la garantía de un elevado *potencial de uso* de los mismos, el cual se logra con una apropiada documentación y estructuración de la información (mediante metadatos y estándares, respectivamente) y con la verificación de contenidos contra fuentes reconocidas como criterios de verdad (esto es, el uso de conjuntos de datos de referencia basados en lenguajes controlados).

Para lo anterior, es importante que quien ejerza las funciones de administrador en el marco del SiB cuente con las siguientes calidades (adaptado de BCIS 2000a y WCMC 1996, 1998a):

- Responsabilidad y honestidad excepcionales, de manera que sea digno de toda confianza.
- Capacidad de liderazgo, enfocada hacia procesos: es muy recomendable que el administrador sepa llevar a cabo procesos que verdaderamente faciliten la gestión de la información que integre al Sistema.
- Capacidad y experiencia en trabajo en equipo.
- Autonomía y criterio para tomar decisiones. El administrador deberá tomar decisiones sobre y la estructuración de sus datos e información y el acceso y uso de la misma, por ejemplo.
- Conocimiento amplio y suficiente de los datos que se generan y manejan en su entorno de trabajo y que estarán bajo su responsabilidad.

Además, debe cumplir dos condiciones mínimas:

- Ser capaz de garantizar la calidad de los datos que administra.
- Conocer, al menos en un nivel básico, herramientas informáticas (*software* y *hardware*). Es deseable que el administrador sea hábil y eficiente en el manejo de las mismas.

V. ¿Cuál es la labor de un administrador en el marco del SiB?

i. Funciones

Un administrador de datos e información tiene dos funciones principales: **mantener** y **custodiar** los datos, la información y los recursos de información que se encuentren a su cargo. **Mantener** significa asegurar que los datos, la información o los recursos de información conserven en el tiempo un elevado potencial de uso y, en esa medida, que puedan ser pertinentes a múltiples procesos en diferentes momentos. Para ello es necesario:

- ✓ **Verificarlos:** revisar que los datos a su cargo sean correctos, consistentes y vigentes, en relación con un punto de comparación o “criterio de verdad”. En el caso de la información sobre biodiversidad, algunos ejemplos de *criterio de verdad* son: un conjunto de datos de referencia (taxonómico, geográfico, metodológico), cartografía digital, un experto (o autoridad) o una publicación científica reconocida, entre otros.
- ✓ **Depurarlos:** seleccionar y excluir aquellos datos que no cumplan con las condiciones arriba mencionadas y que no sean susceptibles de modificación.
- ✓ **Actualizarlos:** realizar las modificaciones que sean del caso para garantizar su calidad y vigencia. Este proceso debe llevarse a cabo periódicamente.
- ✓ Prever su **capacidad de uso e interpretación**, para que a partir de ellos puedan diseñarse productos y servicios de información que respondan a las necesidades de aquellos actores del Sistema que no necesariamente generan información, pero sí la requieren para desarrollar sus propios procesos, tomar decisiones y orientar políticas; estos son los *interesados* y el *público en general* (→ *Productos y servicios*).

Custodiar, por su parte, hace referencia a garantizar la integridad, la seguridad física y la posibilidad de que los datos, la información o los recursos de información sean intercambiados y diseminados en el marco del Sistema. Esto incluye:

- ✓ Garantizar que estén correctamente documentados (→ *Catálogo de metadatos*, → *Registros biológicos*).
- ✓ Asegurar su permanencia en el tiempo, para lo cual es necesario cumplir unas condiciones mínimas en cuanto a su disposición física (→ *Almacenamiento y seguridad física*).
- ✓ Dar a conocer las características o requisitos necesarios para que los datos e información puedan ser compartidos en el marco del Sistema (→ *Definición de condiciones de acceso y uso*).

Debe tenerse en cuenta que no todos los administradores cumplen con todas las funciones anteriormente descritas, puesto que las mismas dependen de cuán involucrados se encuentren con el Sistema y, en último término, de sus propios intereses. Por ello, en el *Esquema general de membresías del SiB* (Franco Villegas 2004) se han establecido dos tipos de administradores:

- ~ **Administradores básicos (administradores de metadatos):** son personas o entidades que documentan con metadatos sus conjuntos de datos sobre biodiversidad, los mantienen, los custodian y declaran su existencia en el dominio público, a través del “Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad”. Las *entidades administradoras básicas* se denominan **nodos administradores de metadatos** y, además de lo anterior, pueden prestar el servicio de compilación y hospedaje de metadatos a otras entidades. Los *administradores básicos* como mínimo custodian los metadatos que generan u hospedan y los verifican y depuran antes de hacerlos públicos.
- ~ **Administradores avanzados (administradores de datos):** son personas o entidades que estructuran sus datos con los estándares generados o recomendados por el SiB, hacen uso de los conjuntos de datos de referencia para documentar o verificar su información en términos taxonómicos, metodológicos o semánticos y también pueden utilizar las aplicaciones de *software* que el SiB ha puesto a disposición de sus miembros para estructurar y publicar su información en Internet. Las *entidades administradoras avanzadas* se denominan **nodos administradores de datos**. Los *administradores avanzados* custodian y mantienen los datos, información o recursos de información.

ii. Derechos

Equivalen a los servicios e insumos que el Sistema pone a disposición de sus socios, como contrapartida, una vez estos se involucran como administradores. La disponibilidad de servicios retribuidos por el SiB está acorde con las responsabilidades propias de cada tipo de administrador, así:

a. Administradores básicos

- ✓ Reciben una capacitación introductoria sobre el esquema SiB para gestión de información sobre biodiversidad y sobre las herramientas técnicas, tecnológicas y de gestión desarrolladas, adoptadas o adaptadas y puestas a disposición por éste.
- ✓ Reciben capacitación sobre el uso y la documentación de metadatos, con el fin de que los metadatos por ellos generados contribuyan al fortalecimiento del “Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad”.
- ✓ Reciben capacitación avanzada sobre el uso de las herramientas puestas a disposición por el SiB para la documentación y consulta en línea de metadatos.
- ✓ Obtienen, siempre que así se establezca, una copia de la aplicación *Cassia*® para documentación de información ambiental, desarrollada en el marco del SiB para la documentación, la administración y la puesta a disposición de metadatos sobre biodiversidad.
- ✓ Son asesorados y apoyados conceptual y metodológicamente en la documentación de sus metadatos y en las actualizaciones periódicas que dicha documentación requiera.
- ✓ Disponen de todos los conjuntos de datos de referencia generados o recomendados por el Sistema (→ *Conjuntos de referencia*).
- ✓ Pueden utilizar la infraestructura del Sistema para poner a disposición de otros sus metadatos (→ *Almacenamiento y seguridad física*).

b. Administradores avanzados

Dado que los administradores avanzados son, a la vez, administradores básicos, reciben todos los servicios e insumos descritos en el apartado anterior y adicionalmente los siguientes:

- ✓ Pueden hacer uso de todas las herramientas técnicas, tecnológicas y de gestión desarrolladas, adoptadas o adaptadas y puestas a disposición por el SiB, entre las que se encuentran:
 - ~ estándares para el manejo de información (→ *Estándares*),
 - ~ conjuntos de datos de referencia (→ *Conjuntos de referencia*),
 - ~ estructuras de datos (→ *Registros biológicos*),
 - ~ estructuras recomendadas de bases de datos (→ *Estructuras recomendadas de bases de datos*),
 - ~ aplicativos (de acuerdo con las condiciones de acceso y uso definidas para ellos),
 - ~ procesos y documentos de gestión, acuerdos y convenios (→ *¿Cómo fluye la información?*).
- ✓ Son asesorados y apoyados conceptual y metodológicamente en la documentación de sus datos e información utilizando formatos estandarizados (→ *Estándares*) y conjuntos de datos de referencia (→ *Conjuntos de referencia*).
- ✓ Reciben orientación en la verificación taxonómica, geográfica, metodológica y semántica de sus conjuntos de datos.
- ✓ Reciben asesoría en el diseño de productos y servicios de información pertinentes a las necesidades del país, que puedan inscribirse en el marco del SiB (→ *Productos y servicios*), a través de las herramientas técnicas y tecnológicas diseñadas por el Sistema.
- ✓ Pueden utilizar la infraestructura del Sistema para poner a disposición de otros sus datos e información (→ *Estructuras recomendadas de bases de datos*, → *Almacenamiento y seguridad física*).

iii. Deberes y responsabilidades

Las responsabilidades y deberes que un administrador tiene con el SiB representan, de acuerdo con la *Política institucional para la gestión de información sobre biodiversidad* (Franco Villegas 2003), compromisos que adquiere al momento de vincularse al Sistema. Los compromisos son extensos, se aplican diferencialmente para cada uno de los tipos de administrador y en todos los casos dependerán de la naturaleza del socio (persona o entidad), del tipo de datos, información o recursos de información que esté administrando; del tiempo de asociación, y de los intereses propios del administrador. Debe aclararse que es posible asociarse al Sistema como administrador sin cumplir la totalidad de los compromisos (en el caso de que la administración sea temporal), siempre y cuando se cumpla el compromiso mínimo:

El deber principal (compromiso mínimo) de vinculación como administrador en el marco del SiB es poner a disposición del Sistema –con libre acceso y uso– y mantener (→ *Funciones*) los metadatos (→ *Metadatos*), en el “Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad” (<http://www.siac.net.co/sib/metadatos/>). La documentación de metadatos deberá realizarse siguiendo los parámetros establecidos por el EC-SiB, con el *Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad* (Rivera-Gutiérrez & Suárez-Mayorga 2005, disponible en línea en http://www.siac.net.co/sib_descargas.php) y mediante la herramienta informática *Cassia*® desarrollada por el Sistema (en línea en: <http://www.siac.net.co/metadatos/>).

Las responsabilidades específicas para cada tipo de administrador están definidas en el *Esquema general de membresías del Sistema de Información sobre Biodiversidad* (Franco Villegas 2004) y se describen en los apartados siguientes.

a. Administradores básicos

- ✓ Deben cumplir el compromiso mínimo arriba mencionado, tanto para los conjuntos de datos nuevos como para los conjuntos de datos antiguos o históricos.
- ✓ Se deben comprometer a mantener actualizados los metadatos disponibles en el “Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad”, haciendo las modificaciones a que haya lugar.
- ✓ En el caso de nodos administradores de metadatos, es su deber alojar los metadatos a las entidades asociadas que así lo requieran y servir de puente de comunicación entre ellas y el EC-SiB.

b. Administradores avanzados

Además de los compromisos propios de los administradores básicos (con excepción del hospedaje de metadatos que únicamente aplica para nodos administradores de metadatos), los administradores avanzados deben:

- ✓ Documentar sus datos de manera compatible con el *Estándar para la documentación de registros biológicos* (Suárez-Mayorga *et al.* 2005a) y estructurarlos para que puedan ser compartidos utilizando el *Estándar para intercambiar información sobre biodiversidad al nivel de organismos* (Suárez-Mayorga *et al.* 2005b).
- ✓ Utilizar los conjuntos de datos de referencia recomendados o puestos a disposición por el Sistema para:
 - ~ documentar correctamente las palabras clave que describan sus conjuntos de datos (→ *Tesaurus*),
 - ~ encontrar información en el Sistema y facilitar que otros socios y usuarios la encuentren (→ *Tesaurus*, → *Entidades y atributos*),
 - ~ verificar (→ *Funciones*) la información taxonómica de sus conjuntos de datos, bien mediante comparación con un archivo de autoridad taxonómica (AAT, → *Archivos de autoridad taxonómica*) o bien, mediante incorporación directa de éste en su estructura de datos,
 - ~ documentar apropiadamente las entidades, los atributos y los métodos que constituyen cada uno de sus conjuntos de datos (→ *Entidades y atributos*).
- ✓ Aportar los atributos y métodos de sus conjuntos de datos al “Catálogo de métodos y atributos” del Sistema (→ *Catálogo de métodos y atributos*).
- ✓ Definir y publicar las condiciones de acceso y uso de los datos contenidos en los conjuntos a su cargo (→ *Definición de condiciones de acceso y uso*) y coordinar con el EC-SiB los mecanismos para garantizar su cumplimiento.
- ✓ Si las condiciones de acceso y uso de sus conjuntos de datos lo permiten, compartir en Internet sus registros biológicos a través del Sistema Distribuido de Búsquedas (SDB) del SiB³, compartir su información sobre especies y ecosistemas en el Catálogo de organismos presentes en Colombia, la herramienta Yoscua para la gestión de información sobre usos y conocimientos tradicionales sobre biodiversidad o cualquier otra herramienta provista o recomendada por el EC-SiB (→ *Productos y servicios*).
- ✓ Poner en práctica los procedimientos de seguridad recomendados por el SiB, con el fin de asegurar la integridad física de los datos. El administrador debe almacenar la información en los formatos y medios necesarios para reducir al mínimo las posibilidades de pérdida, daño, deterioro o acceso no autorizado (→ *Almacenamiento y seguridad física*).

³ http://www.siac.net.co/BancoConocimiento/S/sib_condiciones_acceso_distribuida/sib_condiciones_acceso_distribuida.php

VI. ¿Qué información se administra?

Ángela M. Suárez-Mayorga

i. Datos vs. información

Hasta ahora se ha mencionado que a través del Sistema de Información sobre Biodiversidad se gestionan datos e información para apoyar procesos en torno al conocimiento, la conservación y el uso sostenible de la biodiversidad. Es entonces necesario aclarar que los términos “datos” e “información” no son sinónimos, más bien son insumos secuenciales del “conocimiento”, enmarcados dentro de un sistema particular de pensamiento. O, dicho de otra manera, diferentes niveles de agregación en el proceso de conocer.

Para una buena proporción de los administradores en el marco del SiB, el sistema de pensamiento es el pensamiento científico occidental, pero es importante reconocer que otros sistemas de pensamiento pueden ser (y, en efecto, son) gestionados a través de las herramientas que el SiB proporciona. Este sería el caso de los sistemas que enmarcan el conocimiento tradicional y los saberes locales de comunidades particulares. El reto está, entonces, en integrar estos conocimientos, saberes y sistemas de pensamiento, sin perder de vista el fin común (definido por la visión del Sistema) en el ámbito de la biodiversidad.

En este contexto, los **datos** pueden definirse como:

- ~ observaciones o mediciones de un fenómeno (WCMC 1998a),
- ~ hechos que se registran (son percibidos y documentados) o
- ~ un antecedente necesario para llegar al conocimiento [exacto] de algo o deducir las consecuencias [legítimas] de un hecho (RAE 2001).

Los datos, en el contexto de los sistemas de información, pueden ser tratados desde tres dimensiones: el nivel de interpretación, los análisis que se espera obtener de ellos y la fuente de obtención. Dependiendo del nivel de interpretación, los datos pueden clasificarse así (French *et al.* 1990):

- ~ **Datos crudos:** aquéllos obtenidos directamente de la medición u observación. En el medio científico se suelen almacenar únicamente para propósitos definidos y a corto plazo. Por ejemplo: “20,5°C” en el termómetro húmedo y “19,5°C” en el termómetro seco (son las temperaturas medidas con un higrómetro de campo).
- ~ **Datos calibrados:** valores físicos, corregidos mediante procesos de calibración. Al haber sido sometidos a un proceso de verificación, este tipo de datos usualmente se mantiene o guarda en el tiempo. Ejemplo: 0,9°C (es la diferencia entre los dos termómetros calibrada a la altitud en que fue tomada).

- ~ **Datos validados:** datos calibrados que han sido filtrados por un proceso de aseguramiento de calidad. Son los más usados en procesos de investigación científica. En el ejemplo de arriba, 0,9°C está en el rango de diferencias posibles (0,8 – 1,2°C) entre el termómetro húmedo y el seco de un higrómetro de campo a 1650 m de altitud, de acuerdo con la tabla de conversión del higrómetro.
- ~ **Datos derivados:** datos producto de la agregación de otros datos. Por ejemplo: 0,9°C de diferencia representan una humedad relativa de 91%.
- ~ **Datos interpretados:** datos derivados que se han comparado con otros conjuntos de datos o con literatura. Siguiendo con el ejemplo, puede decirse que la humedad relativa del lugar es muy alta (con relación a parámetros previamente establecidos y ampliamente conocidos). En el contexto más general, *datos interpretados* es equivalente a *información*.

De los datos derivados o interpretados se obtienen los *productos y servicios de información*; por lo tanto, difícilmente son preservados como datos –se preservan en cambio sus contenidos de información, pero en el marco de un análisis para propósitos específicos– (→ *Productos y servicios*). No obstante, la preservación de los datos de todo tipo es vital para cumplir con uno de los objetivos principales del SiB: lograr que los datos que a él se integren puedan ser utilizados para múltiples propósitos y por múltiples usuarios, además de que puedan ser intercambiados fácil y eficientemente (→ *Catálogo de metadatos*). La capacidad de uso de los datos y, por lo tanto, de análisis, es, de hecho, la base del concepto de *calidad* en el Sistema (→ *Calidad y consistencia*) y en el marco de las redes con las que interactúa, como la Infraestructura Mundial de Información sobre Biodiversidad, que en inglés es **Global Biodiversity Information Facility (GBIF)**.

Respecto a la fuente de obtención de los datos, la división tradicional se hace entre datos obtenidos de fuentes primarias y de fuentes secundarias. Las **fuentes primarias** son las observaciones mismas (en el ejemplo la fuente primaria es el higrómetro), mientras que las **fuentes secundarias** representan algún nivel de procesamiento mínimo de los datos que se toma como base (como por ejemplo, la información bibliográfica que permitió establecer que la humedad relativa está catalogada como “muy alta”) para obtener nuevos datos. Dicho de otra manera, los datos de fuentes secundarias se construyen sobre datos ya existentes, usualmente recogidos por una persona diferente (Mayorga Rodríguez 2002).

Partiendo de lo anterior, puede decirse que la **información** es el producto de la interpretación de los datos. Una definición formal del World Conservation Monitoring Centre (WCMC 1998a) dice que la **información** es el resultado de compilar, clasificar, analizar, sintetizar o resumir datos, de forma que estos tengan sentido dentro de un marco de interpretación. Lo anterior implica que la información es una construcción permanente ya que, dado un nivel de agregación o interpretación de los datos, estos siempre serán susceptibles de ser reinterpretados o agregados de diferente manera y constituirán, por tanto, nueva información. En el mismo sentido, la información producto del procesamiento de datos puede ser, a su vez, base (equivalente a datos) para la generación de nueva información. Retomando el ejemplo: los datos indican que la humedad relativa es alta y además, la presión atmosférica es baja; esto es información. La apropiación de esa información por parte de un sujeto (persona o grupo humano) y su aprovechamiento (análisis, transmisión, etc.) constituyen una buena aproximación al concepto de **conocimiento**. La definición de **conocimiento** en el marco de este Sistema está dentro de la línea del concepto social de conocimiento (Polanyi 1958, en Stenmark 2002) e incluye toda la información que se adquiere o acumula o genera, se distribuye o divulga y es apropiada, aprovechada y, a través de esto, validada por un sujeto (individual o colectivo) competente. En el ejemplo, la información mencionada permite al sujeto establecer y anunciar que es conveniente llevar un paraguas al salir pues en breve lloverá.

Sea cual sea la dimensión de clasificación y análisis de los datos y la información, lo más importante en el marco del SiB es que sean de la mejor calidad posible, o, dicho de otra forma, que tengan el más elevado *potencial de uso*. En consecuencia, es necesario proporcionar al usuario todo aquello (metodologías, herramientas, documentación) que le facilite evaluar tal potencial.

ii. Los objetos de interés en el marco del Sistema: entidades y atributos

En términos del SiB, los datos e información se describen y documentan con base en dos conceptos: *entidades y atributos*. Las **entidades** son, de acuerdo con el *Estándar para la documentación de registros biológicos* (Suárez-Mayorga *et al.* 2005a), las *unidades* objeto de estudio, que son descritas mediante *atributos*. Por lo tanto, los **atributos** son las características que describen o explican una entidad; son cualidades propias de la misma. Regresando al ejemplo que se ha venido trabajando, puede decirse que la entidad que se estaba describiendo era el clima del área de estudio y, para ello, se tomaban como atributos la temperatura del aire, la presión, la humedad relativa y la precipitación diaria en el día de la medición.

Entidades y atributos son básicos y suficientes para referirse a los objetos de interés. Sin embargo, es imprescindible articular las entidades con los atributos, que es equivalente a hacer el proceso de descripción de una entidad mediante atributos. Este proceso consiste en definir y aplicar **métodos** que enmarcan la evaluación de los atributos definidos y, en una escala mayor, que definen la captura de información. La definición de **método** en el español es la de “un procedimiento que se sigue en las ciencias para hallar la verdad y enseñarla” (RAE 2001). Aunque en este contexto “la verdad” puede no existir, es claro que un método constituye la conexión necesaria para generar conocimiento a partir de ciertas observaciones interpretadas en un ámbito particular. En términos del ejemplo que se ha venido desarrollando, para decir que el clima (entidad) se describe a partir de los atributos arriba mencionados, es necesario que el investigador asigne valores a los mismos, lo cual se hace con una metodología particular en cada uno de los casos. Para saber la temperatura se utilizará la lectura de un termómetro de mercurio, para hallar la humedad relativa, se utilizará un higrómetro de campo, de forma que el método incluye comparar las temperaturas de un termómetro seco y uno húmedo con respecto a una tabla de corrección; para asignar valores a la precipitación se evaluará el nivel de llenado de un pluviómetro y para hallar la presión, se utilizará la lectura de un barómetro. Nótese aquí que el *método* depende del instrumento de medida.

En el SiB, la mayoría de entidades con las que se trabaja son **unidades biológicas**, es decir, elementos de la biodiversidad en cualquiera de sus niveles de organización o escalas (gen, individuo, población, comunidad, ecosistema) y que se estudian con referencia a tres atributos principales: composición, estructura y funcionamiento (Noss 1990). Sin embargo, como se vio anteriormente, también pueden ser objeto de interés para el Sistema otras entidades: el clima, eventos de uso o de conocimiento tradicional de especies o ecosistemas, o incluso procesos asociados al conocimiento, conservación y uso de la biodiversidad, como el desarrollo de incentivos económicos para conservación.

Teniendo en cuenta la complejidad de la biodiversidad y sus elementos asociados, muchas veces una entidad es, a la vez, un atributo de una entidad mayor. Para ilustrarlo, en la Figura VI-1 se muestra un ejemplo sobre la caracterización de unidades del paisaje con un enfoque jerárquico⁴, donde se represen-

⁴ El ejemplo de la figura se basa en algunos de los métodos para caracterización de la biodiversidad diseñados por la línea de Herramientas de conservación en paisajes rurales, del Instituto Humboldt.

tan en gris las entidades y en blanco los atributos. Sin embargo, cada una de las entidades grises es a la vez atributo de la entidad más oscura inmediatamente superior. De esta forma, la entidad más superior (e inclusiva), que es el paisaje, se describe en términos de otras entidades, como son los elementos del paisaje, la fragmentación y la intervención humana. Cada una de estas entidades es descrita por atributos particulares, como tipo de elemento, distancia entre fragmentos o tipo de uso de la tierra. A su vez, el elemento del paisaje es una entidad que puede ser caracterizada mediante las especies que contiene y la dominancia de alguna de las mismas, mientras que la diversidad de especies puede describirse en términos de riqueza y abundancia.

Para concluir, es importante resaltar que uno de los pilares de la gestión eficiente de información en el marco del SiB es definir acertada y detalladamente las entidades, los atributos y las metodologías empleadas, ya que esto resume y soporta la captura y el intercambio de información (→ *Catálogo de métodos y atributos*), así como la asociación de todos los conjuntos de datos de referencia, los estándares y procedimientos desarrollados en el Sistema.

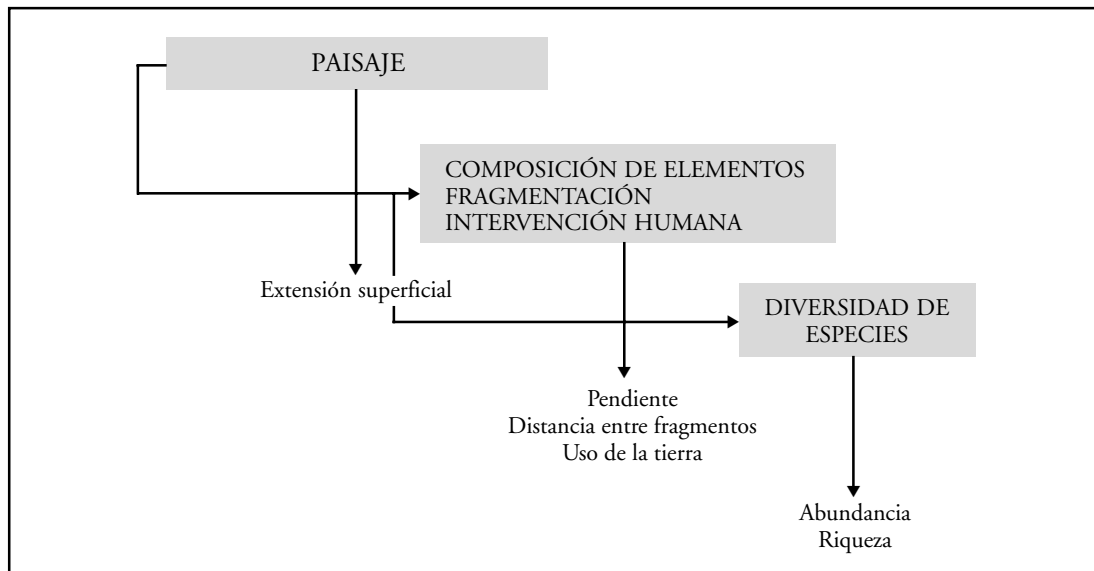


Figura VI-1. Representación de entidades y atributos en la caracterización de la biodiversidad asociada al paisaje.

VII. ¿Cómo se obtiene la información?

Julia Benavides Molineros y Ángela M. Suárez-Mayorga

Por **investigación** se entiende la realización de actividades intelectuales y experimentales de manera sistemática, por parte de un sujeto (que puede ser una colectividad) y con el propósito de avanzar en el conocimiento de un objeto o de determinada materia (Mayorga Rodríguez 2002). Dada el área de conocimiento y los objetos de interés en el Sistema (las unidades biológicas), es un hecho que una buena parte de la información que se vincula al SiB es obtenida a través de observaciones directas o mediante investigación experimental⁵. No obstante, las relaciones de estas unidades con el hombre y la misma necesidad de definir un sujeto que las perciba, dan lugar al uso de otros tipos de investigación, como puede ser la investigación etnográfica⁶ o acción-participativa⁷.

Sea cual sea el tipo de investigación que se realice, el primer paso en la obtención de la información es la formulación de preguntas de investigación para definir, acotar y describir el objeto de estudio y se conoce como **momento lógico** (*ibidem*). Los pasos siguientes (y secuenciales), de acuerdo con la autora arriba mencionada son: el **momento metodológico**, donde se formula un modelo operativo para acercarse al objeto y las estrategias para llevarlo a cabo (definición de metodologías), el **momento técnico**, durante el que se aplican las metodologías para obtener y organizar la información y el **momento teórico**, donde se interpreta la información obtenida en los momentos anteriores. A continuación se hace referencia detallada al **momento técnico**, que es la fase en la que se obtiene la mayoría de los datos e información.

i. La captura de datos

Por el apartado anterior queda claro que la captura de datos debe ser la consecuencia lógica de formular una o varias preguntas de investigación e intentar resolverlas mediante la aplicación de metodologías establecidas que permitan describir las entidades con sus atributos. No obstante, este no es el proceso más frecuente, porque de una pregunta inicial surgen muchas preguntas, de éstas surgen nuevas entidades y

⁵ De acuerdo con Mayorga Rodríguez (2002), la **investigación experimental** es aquella basada en la experimentación, es decir, en comprobar algo por medio de la práctica. Pueden consultarse otras definiciones en Wikipedia (www.es.wikipedia.org) o <http://www.mitecnologico.com/Main/FundamentosDeInvestigacion>.

⁶ Siguiendo a la misma autora, la **investigación etnográfica** tiene por objeto el estudio y descripción de razas, pueblos y grupos humanos, bien desde una óptica externa o bien desde una observación participante que involucra al objeto de estudio.

⁷ **Investigación acción-participativa** hace referencia, según Mayorga Rodríguez (2002), a la participación y el compromiso, a través del intercambio de conocimientos y experiencias entre los sujetos y los objetos.

atributos y se plantean sobre la marcha nuevas metodologías, al punto que en muchas oportunidades se toman datos y datos y al momento de analizarlos no se sabe qué hacer con ellos. Aunque la captura de datos depende siempre de la definición clara de los objetos de interés, hay algunas pautas que es importante tener en cuenta, máxime cuando se piensa en que los datos puedan ser utilizados en diferentes momentos para diversos propósitos:

- ~ Revisión de fuentes. Esto permite, por una parte, verificar que no se esté repitiendo trabajo, de manera que lo que se haga siempre sea un aporte a incrementar el conocimiento que se tiene sobre el objeto de interés o de estudio. Por otra parte, la revisión de fuentes es clave para los análisis de los datos e información capturados.
- ~ Diseño de protocolos de obtención. En la *Política institucional para la gestión e datos e información sobre biodiversidad* (Franco Villegas 2003) se menciona un procedimiento general sobre el ciclo de datos (en su Anexo 1). El primer paso de este procedimiento es la estructuración de los datos previstos (no necesariamente obtenidos) con base en un modelo estándar –el *Estándar para la documentación de registros biológicos* (Suárez-Mayorga et al. 2005a)–, para el caso de las unidades biológicas. Dado que este documento se inscribe en un contexto más amplio, solamente se resaltaré aquí que es necesario predecir el tipo de datos que se obtendrán; así mismo, se deberá definir si habrá un control sobre los valores que cada uno de los atributos (o variables) pueda tomar y qué valores serán permitidos, documentándolos apropiadamente. Para esto el sistema pone a disposición el “Catálogo de métodos y atributos” (→ *Catálogo de métodos y atributos*) y su experiencia en diseñar, si fuera del caso, formatos de captura de información estandarizados que faciliten agregar un componente más de estructura a la colección de los datos mismos. Para ello es necesario:
 - definir qué contenidos requieren uniformidad adicional,
 - estructurar esos campos en un formulario,
 - proveer una lista de valores válidos para dichos campos, y
 - organizar los formularios para que sea fácil utilizarlos en el evento de captura de datos
- ~ Aplicación consistente de las metodologías predefinidas en la obtención de datos: aunque es un hecho que la investigación debe permitir cierto grado de flexibilidad, pues se trabaja con unidades variables en tiempo y espacio, es éste un paso crítico, puesto que de la rigurosidad con que se apliquen las metodologías dependerá la calidad de los datos e información obtenidos.

ii. Calidad y consistencia

En sentido general, **calidad** se define como la “propiedad o conjunto de propiedades inherentes a algo, que permiten juzgar su valor” (RAE 2001) o como el “grado o estándar de excelencia” (WordReference 2004). De acuerdo con las normas ISO 9000, la calidad es todo aquello que una organización hace para cumplir con los requerimientos y las expectativas de sus usuarios (clientes), y para mejorar su desempeño hacia el logro de su misión y sus objetivos (ISO 2007). En el contexto de los datos e información sobre biodiversidad, BCIS (2000b) propone que “la calidad de los datos es una medida de ‘aptitud para el uso’ y por consiguiente sólo puede establecerse con respecto a un uso esperado. Los datos pueden ser de alta calidad para un propósito, pero de baja calidad para otro. Si bien es importante que los datos tengan la mayor calidad posible, es más importante que se documente dicha calidad, para permitir que un usuario cuente con la información suficiente para decidir si esos datos corresponden o no al uso esperado, en lugar de pugnar por un ideal de calidad teórico”.

En el Sistema de Información sobre Biodiversidad pueden identificarse varios niveles de calificación de la calidad de los datos e información:

- ~ El autor o custodio de conjuntos de datos (o quien documente los metadatos de esos conjuntos) debe registrar la precisión de los atributos medibles.
- ~ El administrador de conjuntos de datos de referencia debe calificar y validar los datos que incluye en el conjunto de datos respectivo.
- ~ El administrador de conjuntos de datos distribuidos suministra información objetiva al usuario para que éste califique los conjuntos de datos.

El **usuario** de los datos debe tener herramientas o información suficiente para poder formar su propio juicio de valor sobre los datos disponibles y poder juzgar la pertinencia de usarlos con un propósito específico.

Debe anotarse con toda claridad que la confiabilidad y exactitud de los datos de conjuntos de datos distribuidos que sean integrados al Sistema de Información sobre Biodiversidad son responsabilidad de su autor o autores. Ni el EC-SiB, ni el Sistema en su conjunto, ni el Instituto Humboldt, son responsables por el uso que se pueda hacer de estos datos, debido a que ninguno de estos entes tiene control sobre su origen, ni poder de validación (adaptado de REMIB⁸).

iii. Herramientas de evaluación de calidad

a. Estructuras y conjuntos de datos centrales

El Sistema de Información sobre Biodiversidad ha desarrollado varias herramientas para soportar la calidad de sus contenidos, expresados en los productos y servicios de información (→ *Productos y servicios*). Estas herramientas son de tres tipos: las estructuras y conjuntos de datos centrales, los documentos de la estructura del sistema y la lista de criterios de evaluación de la calidad. A continuación se presenta la relación entre estas herramientas y su papel en el aseguramiento de la calidad de los datos (Tabla VII-1).

b. Documentos de la estructura del Sistema

Estos documentos abordan tres aspectos de las herramientas y la estructura general del sistema: **cómo se hace, qué contiene y cómo funciona/cómo se usa**. El primer aspecto generalmente está contenido en los protocolos, el segundo en los estándares y el tercero en las guías de usuario. En la Tabla VII-2 se presenta un resumen de estos documentos.

⁸ URL <http://www.conabio.gob.mx/remib>

Tabla VII-1. Descripción de las herramientas del sistema que contribuyen a respaldar la calidad de sus contenidos.

Herramienta	Cómo contribuye a garantizar la calidad
Estructura de registros biológicos	Garantiza la integridad a los datos. Permite la interoperabilidad de los datos (aunque debe anotarse que la interoperabilidad no es condición <i>sine qua non</i> de la calidad).
“Catálogo de métodos y atributos”	Suministra a los investigadores información sobre las diversas metodologías existentes para abordar problemas similares. El investigador puede seleccionar aquellas metodologías que permitan que sus datos sean comparables.
AAT	Sirven para evitar errores en la escritura de los nombres. Garantizan el uso de nombres correctos (sirve para evitar el uso de sinonimias).
“Tesauruso sobre biodiversidad”	Garantizan la calidad en la asignación de descriptores, en el sentido de que éstos representen adecuadamente el objeto que describen. Permiten obtener resultados exitosos en búsquedas sobre un tema específico mediante el uso de las relaciones de parentesco entre los términos.
“Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad”	Suministra al usuario información sobre autor, propósito, cobertura geográfica, métodos, atributos y condiciones de acceso de los conjuntos de datos. Con esta información el usuario puede formar su juicio de valor sobre los datos.
Aplicación <i>Cassia</i> ®	Permite a los usuarios describir sus conjuntos de datos en formatos estándar (bien sea el estándar nacional u otros estándares acordados previamente), y facilita la publicación en Internet de los metadatos documentados, haciendo uso del “Catálogo nacional de metadatos de conjuntos de datos relacionados con biodiversidad”.

Tabla VII-2. Descripción de documentos del sistema que contribuyen a respaldar la calidad de sus contenidos. Todos los que se han resaltado en cursiva se encuentran disponibles en la sección de Descargas del portal Web SiB (http://www.siac.net.co/sib_descargas.php).

Cómo se hace (protocolos)	Qué contiene (estándares)	Cómo funciona, cómo se usa (guías de usuario)
Implementación de catálogo de metadatos. Documentación de metadatos	<i>Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad.</i>	
	<i>Estándar para la elaboración de archivos de autoridad taxonómica.</i>	<i>Manual de usuario de la aplicación de AAT.</i>
Documentación de registros biológicos.	<i>Estándar para la documentación de registros biológicos.</i> <i>Estándar para intercambiar información sobre biodiversidad al nivel de organismos.</i>	<i>Guía para integrar conjuntos de datos de referencia usando Microsoft Excel®</i> <i>Guía para integrar conjuntos de datos de referencia usando Microsoft Access®</i>
<i>Propuesta para la implementación de los tesaurusos sobre biodiversidad en el IAvH.</i> <i>Protocolo para la elaboración del tesauruso del IAvH sobre biodiversidad.</i>		<i>Guía para la elaboración de tesaurusos monolingües.</i>

(Continúa)

(Continuación)

Cómo se hace (protocolos)	Qué contiene (estándares)	Cómo funciona, cómo se usa (guías de usuario)
<i>Metodología para la clasificación formal de términos del tesoro del IAvH sobre biodiversidad.</i>		
	<i>Documento explicativo del catálogo de métodos y atributos.</i>	<i>Documento explicativo del catálogo de métodos y atributos.</i>
Documentación de aplicaciones	<i>Manual de la aplicación de registros biológicos.</i>	<i>Manual de la aplicación de registros biológicos. Guía para la evaluación de software para la administración de colecciones biológicas.</i>

c. Criterios de evaluación de calidad

La evaluación del nivel superior incluye los criterios de evaluación de los niveles inferiores; por ejemplo, para evaluar la información, también se aplican los criterios de evaluación de datos y conjuntos de datos.

- *Criterios de evaluación de datos y conjuntos de datos*
 - Origen de los datos
 - Métodos empleados en la recolección de datos
 - Exactitud en atributos medibles
 - Precisión en atributos medibles
 - Estructura lógica
 - **Consistencia dentro del conjunto:** coherencia entre los elementos del conjunto
 - **Complejidad de los registros:** medida de la cantidad de información almacenada en un conjunto de datos en relación con el potencial de información que es posible almacenar en él.
 - **Control de procesos relacionados con el almacenamiento de datos:** control de eliminación y modificación de registros, rastreo de cambios, reglas de validación que impidan que se ingresen registros incompletos, con combinaciones erróneas de atributos, errores tipográficos o registros duplicados.
 - **Eficacia en el almacenamiento:** integridad y seguridad.
 - **Eficacia en la recuperación de los datos:** uso de tesauros, documentación de los conjuntos con metadatos.
 - Mantenimiento y actualización de los datos.
 - **Potencial de integración de los datos:** estructura del conjunto, interoperabilidad con otros conjuntos.

- *Criterios para la evaluación de información*
 - Coherencia conceptual entre problema de investigación y los datos y métodos usados para abordarlo.
 - Métodos usados para el procesamiento y análisis de los datos.
 - Objetividad de los métodos de análisis.
 - Disponibilidad de la información generada.
 - Utilidad de la información para la toma de decisiones.

VIII. ¿Qué significa “datos administrados en el marco del SiB”?

Ángela M. Suárez-Mayorga, Héctor Fabio Rivera-Gutiérrez, Talía Waldrón y Diego Trujillo-Motta

A lo largo de este documento se ha mencionado muchas veces que los datos e información se administran en el marco del SiB. Como se ha visto, esto tiene varias implicaciones: su calidad y consistencia deben garantizarse, deben poder ser intercambiados eficientemente y deben poderse utilizar en diferentes momentos y para múltiples propósitos (entre otros). Sobre el tema de la calidad se habló en secciones anteriores, así como sobre los métodos, entidades y atributos; también se mencionaron algunas condiciones para un manejo óptimo de los datos entre las funciones, deberes y derechos de los administradores. Sin embargo, en este capítulo se detalla el tratamiento que debe darse a los mismos para convertirlos en interoperables y multipropósito.

i. Estandarización: datos interoperables

Cuando se trabaja con un esquema de información distribuida como el que propone el Sistema de Información sobre Biodiversidad (SiB), las herramientas para lograr un intercambio y diseminación eficiente de información se basan en la unificación de los contenidos, con el fin de hacer interoperables e intercambiables los conjuntos de datos integrados al Sistema. De acuerdo con Weibel (1998) existen tres niveles de interoperabilidad: semántico, sintáctico y estructural. La interoperabilidad semántica (de significados) se basa en la generación de acuerdos sobre los contenidos del sistema, mediante el uso de lenguajes controlados; el objetivo es ponerse de acuerdo en lo que se dice, para que aunque se usen palabras diferentes, tengan el mismo sentido cuando son utilizadas en contextos similares. La interoperabilidad estructural está dada por la manera en que los contenidos se organizan y almacenan; en los sistemas de información se expresa en modelos de datos que permiten definir e incorporar esquemas semánticos que puedan ser compartidos. Finalmente, la interoperabilidad sintáctica se refiere a la manera en que los datos se ordenan y son “marcados” de manera similar para cada conjunto y entre conjuntos, permitiendo que las máquinas (computadores) entiendan la forma en que deben procesarse y transferirse con el objeto de ser compartidos e intercambiados. Este último tipo de interoperabilidad está soportado, por ejemplo, por un lenguaje llamado eXtensible Markup Language (XML).

a. Estándares

En la escala de la interoperabilidad estructural, el Sistema ha diseñado herramientas desde la administración de contenidos que hacen posible mantener una uniformidad conceptual (semántica) entre los datos

e información. Dentro de estas herramientas se cuentan los diferentes estándares, los protocolos y los procedimientos.

De manera formal, para el SiB un **estándar** es un documento que sirve como herramienta de patrón o referencia para la unificación de información y del formato en el cual esta información es consignada, registrada o intercambiada (Rivera-Gutiérrez *et al.* 2003b). Partiendo de esta definición, todos los estándares puestos a disposición por el Sistema hacen referencia a dos temas principales: qué información se debe registrar y de qué manera se registra.

La información que se debe registrar se representa en el documento de estándar como **elementos de datos**. De acuerdo con su tipo, los elementos de datos pueden ser simples o compuestos (definidos en sí mismos o definidos por otros elementos de datos, respectivamente), pero es en los simples donde se ingresa información. De acuerdo con su obligatoriedad, los elementos de datos pueden ser *obligatorios* (como por ejemplo el contacto responsable de la identificación de un ejemplar de colección), *obligatorios si aplican* (sólo se tienen en cuenta si es posible aplicarlos) y por último, opcionales. En un marco más general, puede decirse que los elementos de datos hacen referencia a conceptos específicos y son equivalentes a los campos de una base de datos que se deben completar.

La manera de registrar la información se presenta en el documento de estándar como **formas de representación**. Las formas de representación son especialmente importantes a la hora de estructurar una base de datos, pues de ellas dependerán la longitud de los campos y los valores válidos, entre otros. Pero primordialmente son importantes para garantizar una comunicación eficiente entre conjuntos de datos –su interoperabilidad–. Los estándares (como es el caso de los propuestos por el SiB) pueden contener también **secciones recurrentes**, que son numerales compuestos por varios elementos de datos a los que se hace referencia varias veces en el curso del documento y cuya composición y estructura no varía entre dichas referencias; de hecho, estos parámetros son iguales en todos los estándares del SiB. Ejemplos de secciones recurrentes son la *información de citación*, la *información de contacto* y las metodologías.

Actualmente, el SiB cuenta con cuatro estándares desarrollados y adaptados para facilitar la interoperabilidad estructural entre conjuntos de datos (sea cual sea su origen y lugar de almacenamiento):

- ~ *Estándar para la documentación de registros biológicos* (Suárez-Mayorga *et al.* 2005a). Es la base de la estructura conceptual y física sobre la que se desarrolla el Sistema. El propósito del documento es recomendar el conjunto de elementos de datos mínimos necesarios para documentar apropiadamente cualquier captura de información sobre una unidad biológica. La estructura y el contenido de este estándar se basan en el estándar internacional de metadatos con perfil biológico de la National Biological Information Infrastructure (NBII) de Estados Unidos y son resultado del trabajo del equipo coordinador del SiB con los diferentes actores involucrados en el proceso de gestión de información biológica. Para garantizar interoperabilidad, se siguen las recomendaciones para bases de datos de colecciones biológicas elaboradas por Berendsohn *et al.* (1999).
- ~ *Estándar para intercambiar información sobre biodiversidad al nivel de organismos* (Suárez-Mayorga *et al.* 2005b). Este estándar comprende la información que es imprescindible poner a disposición a través de Internet cuando se describe la presencia de organismos en un lugar determinado, bien sea que éstos formen parte de una colección biológica o que hayan sido registrados por cualquier medio (observaciones, grabaciones, imágenes, entre otros). A diferencia del *Estándar*

para la documentación de registros biológicos, el estándar de intercambio no regula la forma en la que se captura la información, sino los contenidos que es necesario compartir para poder integrar la información proveniente de cualquier estructura de datos. Por ello, su compatibilidad con otros estándares desarrollados para tal fin en el mundo debe estar garantizada. En este sentido, el referente principal es el estándar Darwin Core (versiones 1.4 y 2, así como la versión actual Wiki⁹) que es el estándar para documentación e intercambio de información biológica más ampliamente utilizado.

- ~ *Estándar para la elaboración de archivos de autoridad taxonómica* (Trujillo-Motta *et al.* 2005). En él se establecen y definen cada uno de los elementos de datos requeridos para la elaboración de un archivo de autoridad taxonómica (AAT), puesto que aunque recientemente el concepto de AAT ha sido incorporado en diferentes iniciativas de gestión de información sobre biodiversidad en el mundo (la Conabio¹⁰ de México, NBN¹¹ del Reino Unido y NBII¹² de Estados Unidos, entre otras), ninguna de éstas ha formulado un protocolo que defina en forma estructurada cada uno de los elementos que deben conformarlo. No obstante, el grupo de trabajo de bases de datos sobre biodiversidad (TDWG) está fomentando el uso de una estructura concertada, el Taxonomic Concept Schema, para documentar la información correspondiente a nomenclatura taxonómica. Para garantizar la interoperabilidad con otros conjuntos de referencia similares, este estándar fue elaborado teniendo en cuenta algunas propuestas desarrolladas previamente por Bisby (1994), Berendsohn (1997) e ITIS (2001). Frente a ellas, el estándar de AAT tiene características que lo hacen una propuesta muy sólida, como que cada uno de los atributos definido está sustentado por sus respectivas referencias bibliográficas y que su concepción está basada en las disposiciones establecidas por los códigos de nomenclatura botánico (Greuter *et al.* 2000) y zoológico (ICZN 2000), por lo cual es aplicable para organismos considerados en cualquiera de los reinos Animalia, Plantae o Fungi; además contiene elementos que permiten la documentación del sistema o sistemas de clasificación considerado(s) en la elaboración de un AAT.
- ~ *Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad* (Rivera-Gutiérrez y Suárez-Mayorga 2005). Este documento es una recomendación técnica que define el conjunto y tipo de datos mínimos requeridos para que las descripciones sobre cualquier conjunto de datos que tenga relación con biodiversidad al nivel nacional sean las más adecuadas y respondan a los mismos criterios. Por lo tanto, constituye una guía para establecer cuáles son los elementos que se pueden y deben utilizar en la descripción de los diferentes conjuntos de datos que se integran al Sistema y esto es equivalente a **documentar apropiadamente** (describir por medio de documentos, con unos contenidos estandarizados mínimos) tales conjuntos de datos a partir de *metadatos* (→ *Catálogo de metadatos*). El estándar ayuda a identificar los elementos que describen adecuadamente cada conjunto de datos, proporciona indicaciones sobre los posibles valores que puede tomar cada uno de los elementos y define cada uno de los conceptos que permitirán construir la descripción de un conjunto. Para ello está dividido en seis secciones, representando los principales aspectos del conjunto de datos: identificación, coberturas, responsable de la información, referencias cruzadas, distribución de

⁹ Actualmente Darwin Core continúa desarrollándose como una iniciativa bajo el liderazgo de John Wiecezorek en el marco del Grupo de trabajo de estándares sobre biodiversidad, *Taxonomic Databases Working Group* (TDWG, <http://www.tdwg.org/>). Las diferentes versiones de Darwin Core pueden consultarse en línea en <http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome>

¹⁰ Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (<http://www.conabio.gob.mx/>)

¹¹ National Biodiversity Network (<http://www.nbn.org.uk/>)

¹² National Biological Information Infrastructure (<http://www.nbi.gov/>)

la información y referencia de metadatos. Para facilitar la utilización de este estándar en la documentación de metadatos, el Equipo Coordinador del SiB desarrolló el Catalogador SiB sobre información ambiental (*Cassia*®), que es una aplicación informática en línea para documentar y hacer disponibles en Internet metadatos de todo tipo. Solamente es necesario definir y seleccionar un estándar de documentación y solicitar una contraseña de usuario (→ *Catálogo de metadatos*).

En cualquiera de estos estándares se explica detalladamente cada uno de los componentes arriba mencionados y su representación.

Es recomendable que el administrador conozca y entienda cada uno de los estándares adoptados por el Sistema, puesto que serán básicos para su trabajo como miembro del mismo.

b. Conjuntos de referencia

En el apartado anterior se habló de los niveles de interoperabilidad, haciendo especial referencia a la interoperabilidad estructural. En este apartado se tratará el nivel de interoperabilidad semántico, que significa tener contenidos unificados mediante **lenguajes controlados**. Un **lenguaje controlado** es un sistema de conceptos y símbolos que se estructura mediante procedimientos de control de vocabulario. En la gestión de información, un lenguaje natural controlado es necesario y útil porque facilita la recuperación de información sobre un tema específico y asegura al usuario encontrar todos los posibles contenidos relacionados con un concepto dado, siempre que cada uno esté asociado con un término, o vinculado con un grupo de términos equivalentes. Por esta razón, una de las características de este esquema es el manejo de **conjuntos centralizados de referencia** –administrados por una sola organización que, en este caso, es el Equipo Coordinador del SiB, en el Instituto Humboldt–, los cuales facilitan la interoperabilidad y la integración de los diferentes recursos de información sobre biodiversidad. Estos conjuntos centrales de referencia son: los *archivos de autoridad taxonómica* para los grupos de organismos presentes en el país o incluso, en el Neotrópico; el *catálogo de métodos y atributos* y el *tesauro sobre biodiversidad*. Los tres, enmarcados en el SiB, constituyen lenguajes controlados vitales para la estructura conceptual y tecnológica que el SiB ha desarrollado y durante la búsqueda y recuperación de la información operan como un puente de comunicación entre las bases de datos que conforman un sistema distribuido.

En este punto vale anotar que en el estudio de la naturaleza existen diferentes aproximaciones y la que prima en la actualidad es la que invita a discriminar unidades biológicas que, aunque varían de acuerdo con su complejidad (gen-especie-población-comunidad-ecosistema-paisaje-bioma) pueden considerarse unidades discretas, y evaluarse con tres conjuntos independientes de atributos (Gaston 1996): composición (identidad y variedad de elementos), estructura (organización física o patrones de los elementos) y función (procesos ecológicos y evolutivos) (→ *Los objetos que se gestionan en el marco del Sistema*).

Si se parte de una idea de conocimiento de la naturaleza en la cual los organismos o las entidades aparecen en el momento que son nombradas y además se asume que las especies no cambian su apariencia física, puede tenerse una idea de cómo pensaba Linneo y de qué manera se enfrentaba con la explosión de descripciones de especies ocasionada por la expansión colonial. Es así como Linneo encontró la forma de

categorizar jerárquicamente los organismos de acuerdo con criterios anatómicos y fisiológicos (ignoró los ecológicos) y propuso como clases más generales tres reinos: Plantas, Animales y Minerales, y algunas de las categorías que aún se mantienen: reino, phylum, clase, orden, familia, género y especie, donde cada una de estas clases corresponde a un concepto determinado. Linneo también aplicó el uso de una nomenclatura binomial en latín como **nombre científico** para las diferentes especies de organismos, lo que ha permitido el intercambio de información precisa entre los especialistas. Actualmente existen *archivos de autoridad taxonómica (AAT)* elaborados por los especialistas de cada grupo de organismos (p.e. peces, mamíferos) que representan un vocabulario controlado necesario para el intercambio de información biológica. Además, la existencia de nombres científicos normalizados permite buscar sus equivalencias (si las hay) en los nombres comunes de organismos usados por diferentes comunidades locales (p.e. indígenas, afrodescendientes, campesinos), ofreciendo otro producto de información sobre la naturaleza basado en el uso de los términos.

Todas las ciencias tienen un lenguaje particular que permite la comunicación entre especialistas y si se hallan las equivalencias con el lenguaje natural, ese mismo lenguaje permitirá al público en general acceder a información sobre biodiversidad que pueda interesarle, como por ejemplo ¿cuáles organismos sólo se encuentran en Colombia?, ¿cómo se usan las palmas de vino en la zona andina?, ¿qué es una especie invasora?, ¿cómo se determina si una especie está amenazada?, ¿cuál es el cultivo que ocupa mayor área en el territorio colombiano?, ¿cómo afectan la fumigaciones aéreas a las comunidades que viven expuestas a ellas?, ¿quiénes usan la cruz de mayo en Colombia y cuál es la planta que usan para hacerla?, ¿cuál es la flor más grande que existe?, ¿cuánto dura la vida de un escarabajo?, ¿cuáles plantas deben sembrarse juntas para que crezcan en las mejores condiciones?

En los apartados siguientes se describirá detalladamente cada uno de los conjuntos de referencia construidos y puestos a disposición por el SiB.

- Archivos de autoridad taxonómica

En la administración de datos biológicos, el papel de la taxonomía va más allá de la comunicación básica. Los nombres científicos son elementos clave usados para la recuperación, organización e intercambio de información biológica (Blum 1998). Es por esto que el control de las inconsistencias nomenclaturales en las bases de datos biológicas es un proceso fundamental para asegurar la integración de los conjuntos de datos sobre biodiversidad. Este control se logra a través del uso de **archivos de autoridad taxonómica (AAT)**, que son lenguajes controlados que facilitan el *control de calidad* de la información nomenclatural básica que se ingresa a las bases de datos biológicas (Escalante *et al.* 2000) y contribuyen a garantizar la integridad nomenclatural en las base de datos. De acuerdo con esta definición, los AAT son conjuntos de referencia que deben administrarse en forma centralizada, por lo cual en la propuesta de implementación para Colombia del Sistema de Información sobre Biodiversidad se determinó que la administración de este conjunto esté a cargo del equipo coordinador del Sistema (EC-SiB), que en su estructura organizacional incluye un administrador de AAT encargado de las labores correspondientes a la gestión de este conjunto de datos.

A pesar de lo anterior, en la mayoría de los casos la integración de estos archivos previa a la digitalización de los datos o la posterior verificación de los nombres científicos contra los mismos requieren que tanto los nombres como sus datos nomenclaturales básicos (los autores y años, entre otros) estén estructurados en forma adecuada.

La estructura de los nombres científicos de los archivos de autoridad taxonómica elaborados en el marco del SiB es recursiva; esto indica que un nombre completo (género más epíteto) siempre es construido con referencia al nombre del taxón superior o taxón padre en la jerarquía taxonómica. Por lo tanto, se recomienda que los nombres taxonómicos en las bases de datos que serán integradas al SiB siempre vayan en campos separados uno seguido del otro, como se muestra en el siguiente ejemplo (Tabla VIII-1):

Tabla VIII-1. Ejemplo de estructuración de nombres científicos en el marco del SiB.

Familia	Género	Epíteto específico	Epíteto subespecífico	Autor
Cerambycidae	<i>Lochmaeocles</i>	<i>tessellatus</i>	<i>tessellatus</i>	(Thomson, 1868)
Orchidaceae	<i>Masdevallia</i>	<i>tubulosa</i>	<i>syringodes</i>	(Luer & Andreetta) Luer

Los autores (y años en el caso de la nomenclatura zoológica), se requieren para diferenciar los homónimos, es decir, aquellos nombres idénticos asociados a taxones diferentes. Por ejemplo, *Potamophylax* Wallegren, 1891 y *Potamophylax* Myers Carvalho, 1955; el primer género fue descrito para un insecto y el segundo, para un pez.

La recomendación del EC-SiB es que los AAT o recursos taxonómicos similares sean descargados e integrados en las bases de datos antes de la sistematización de los registros biológicos. Sin embargo, los nombres científicos y su información nomenclatural y taxonómica asociada que ya haya sido sistematizada, podrán ser verificados también con respecto a los AAT disponibles, para corroborar su validez, evitar o reducir los errores tipográficos y asegurar el uso de datos fiables y actualizados. Para ello, existen herramientas disponibles en línea como la aplicación *Salvias Taxon Scrubber* (Boyle 2004, <http://www.salvias.net/pages/taxonscrubber.html>) y las dos guías de usuario para la integración de conjuntos de datos de referencia –utilizando Microsoft Excel© y Microsoft Access©– que el EC-SiB ha generado, y que se encuentran disponibles impresas y en línea (http://www.siac.net.co/sib_descargas.php, dentro de la carpeta “Registros biológicos” y la subcarpeta “Documentos”). En el apartado siguiente se describen los procedimientos para la consulta y descarga de los AAT (→ ¿Qué significa “datos administrados en el marco del SiB”?, *Procedimientos asociados*).

- Tesaurus

El **tesauro** en el marco del SiB se define como el conjunto de términos que describe todos los conceptos relacionados con un área de conocimiento. En este caso el área de conocimiento es la biodiversidad presente en el territorio colombiano, de tal manera que incluye los términos biológicos, políticos y culturales.

Un tesauro permite describir un campo de estudio mediante un conjunto de términos normalizados, como se verá a continuación. Partiendo de un término cualquiera, y sin ayuda de su significado –sólo con base en las relaciones existentes entre términos–, se puede aclarar éste y sugerir todos los posibles términos relacionados con él. Estas relaciones existentes entre términos se denominan **jerárquicas** (un término está contenido en otro de acuerdo con su significado), de **equivalencia** (con términos sinónimos

o cuasi-sinónimos) y **asociativas** (con términos relacionados). Además, el tesoro puede guiar la búsqueda de un tema específico usando las relaciones de parentesco entre los términos, de modo que la información obtenida al estudiar y comparar dichas relaciones responda a la consulta definida por el usuario. Lo anterior permite, entre otros, formar puentes hacia todos los aspectos relacionados con la biodiversidad: su conocimiento, uso y conservación.

En el Tesoro sobre biodiversidad del SiB, es posible también describir apropiadamente conjuntos de datos biológicos que se gestionan en el sistema y buscar información al interior del mismo, garantizando que se seleccionen descriptores ampliamente utilizados y, por tanto, seleccionados con frecuencia en las búsquedas. En el apartado iii de este capítulo se explican los procedimientos necesarios para seleccionar descriptores o realizar búsquedas orientadas mediante tesauros.

- Catálogo de métodos y atributos

El catálogo de métodos y atributos surge como una derivación del Catálogo de metadatos, con el objeto de describir los atributos considerados en los conjuntos de datos integrados al sistema y los métodos empleados para evaluar esos atributos o para obtener los datos e información que constituyen los registros. Este catálogo es aproximadamente equivalente a un diccionario de datos (muy frecuente en aplicaciones SIG), puesto que describe aspectos conceptuales y estructurales de los datos: los conceptos contemplados, las definiciones, los valores que puede tomar cada concepto (**dominios**) y los tipos de valores (**tipos de dominio**) a los que corresponden los valores seleccionados, y que están siendo representados en un conjunto de datos particular para una estructura dada. Sin embargo, existen diferencias sustanciales:

- ~ Todas las definiciones de métodos y atributos están relacionadas con una citación fuente y/o un contacto responsable, los cuales pueden ser consultados en los repositorios de citaciones y contactos de las bases de datos.
- ~ El catálogo tiene una estructura jerárquica que permite asociar los conceptos expresados por los métodos y atributos a través de los términos que los representan. Por lo tanto, tiene las mismas propiedades de un tesoro en cuanto a lenguaje controlado. Adicionalmente, todos los métodos y atributos son identificados en el sistema por palabras clave, que los vinculan directamente al tesoro general y permiten la conexión entre los diferentes conjuntos de datos integrados.
- ~ Se pueden identificar fácilmente (y describir) las relaciones entre métodos y atributos, es decir, qué atributos están referidos a un método particular y qué métodos consideran unos atributos particulares en su definición. Dado que las relaciones entre métodos y atributos son múltiples (el mismo método puede estar caracterizado por diferentes atributos o el mismo atributo puede formar parte de diferentes métodos), el catálogo las hace explícitas a través de la interfaz de consulta.

Por todo lo anterior, puede decirse que el Catálogo de métodos y atributos es parte de la memoria documental de cualquier conjunto de datos, pero también la principal herramienta para evaluar la calidad de los contenidos de éste. Los métodos y atributos correctamente documentados proporcionan una referencia rápida a los contenidos de los conjuntos de datos en el sistema, detallando las características con las que fueron definidos inicialmente. Por lo tanto, al consultar el catálogo es posible obtener una idea de la utilidad y pertinencia de los conjuntos de datos y, por ende, su calidad (→ *Calidad y consisten-*

cia). Por último, al recopilar con detalle características propias de los atributos y métodos, puede ser la base de estructuras de bases de datos flexibles que permitan al usuario seleccionar los atributos a considerar en el conjunto de datos particular, de acuerdo con sus necesidades e intereses.

Actualmente, el catálogo recoge un detalle de la información que se debe documentar en las secciones 5 y 6 del *Estándar para la documentación de registros biológicos* (Suárez-Mayorga *et al.* 2005a), a excepción de los valores propios que toma cada atributo en cada registro particular. Por lo tanto, para cada método o atributo documentado en el catálogo, se encontrará: 1) nombre, 2) definición o descripción, 3) fuente de la definición y 4) contacto de la definición, mientras que para los atributos se describen además el dominio y el tipo de dominio, que se definen así:

- ~ *tipo de dominio*: se refiere al tipo de valores que puede tomar el atributo. Puede ser de seis clases: 1) dominio enumerado, cuando los valores posibles pueden preestablecerse en su totalidad, se mencionan sucesivamente y son mutuamente excluyentes, como en el caso de los colores primarios; 2) dominio múltiple, cuando los valores posibles pueden preestablecerse en su totalidad, se mencionan simultáneamente y no son mutuamente excluyentes, 3) dominio de código, cuando los valores pertenecen a un código formalmente registrado y publicado, y así se enumeran, como por ejemplo las categorías de amenaza de la UICN (2001); 4) dominio de rango, para los valores numéricos que forman parte de los número reales y que usualmente evalúan atributos producto de mediciones, como es el caso de la estatura (valores entre 40 y 250 cm); 5) dominio compuesto, cuando los valores que lo componen son, a su vez, atributos individuales con su propio dominio; es el tipo de dominio que se utiliza para detallar categorías de uso. Por último, 6) dominio no representable, que se utiliza cuando no hay manera de predefinir los valores que el atributo puede tomar, por lo que hay que exponer las razones que llevan a esta condición. El ejemplo más típico es el de los nombres comunes para organismos;
- ~ *dominio*: es el conjunto detallado de todos los valores que el atributo puede tomar o de las razones por las que es imposible prever dichos valores y documentarlos, en el caso de dominios no representables. Siguiendo con los ejemplo anteriores, para el dominio enumerado los valores serían “amarillo”, “azul”, “rojo”; para el dominio múltiple los valores pueden ser “amarillo y azul”, “amarillo y rojo”, “azul y rojo”, “amarillo”, “azul” o “rojo”; para el dominio de código los valores son “EX”, “CR”, “EN”, “VU”, “DD”, “NE” y para el dominio de rango los valores responderían a la expresión $40 \leq \text{Estatura (reales)} \leq 250$. En el caso de dominios compuestos, como el del atributo “Categoría de uso”, los valores que podría tomar son “uso medicinal”, “uso ornamental” y “uso como transporte”; cada uno de esos valores es, a su vez, un atributo con dominio múltiple que describe todos los usos medicinales, ornamentales y como transporte posibles. Finalmente, para el dominio no representable el valor sería cualquier texto, ya que los nombres pueden ser tan diversos como la creatividad de las poblaciones locales en diferentes localidades.

Adicionalmente, para los atributos se registran varios elementos adicionales. Los dos primeros de la lista que se menciona a continuación se registran en todos los casos, mientras que los dos últimos se toman únicamente para atributos medibles:

- ~ **valor de exactitud**: es una estimación hecha mediante una descripción o un porcentaje, de la confianza con la que se asignan valores a un atributo particular. Por ejemplo, si en un conjunto de datos de aves se define el atributo “número de anillo” para las aves migratorias marcadas y la

identificación de este número solamente es posible teniendo el pájaro en la mano, el valor de exactitud podría ser del 95 - 99%, pues la única posibilidad de que el valor registrado difiera del valor real es una mala apreciación del observador, pero difícilmente se equivocará a menos que sufra problemas de hipermetropía;

- ~ **frecuencia de actualización:** es un indicador (numérico o descriptivo) de la frecuencia con la que se ingresan valores a un atributo. En el ejemplo anterior, “número de anillo” solamente tiene nuevos valores cuando se captura un ave migratoria marcada;
- ~ **unidad de medida:** indica las unidades a las que hace referencia la medición de un valor para el atributo;
- ~ **precisión:** permite registrar la precisión del aparato de medida, como por ejemplo 0.1 cm en el caso de la estatura.

c. Procedimientos asociados

⇒ Consultas de AAT

El usuario que sólo esté interesado en consultar los datos taxonómicos o nomenclaturales relacionados con un taxón, podrá hacer su consulta a través de la página de Administración de contenidos del SiB¹³ o directamente en la página de Archivos de Autoridad Taxonómica¹⁴. Para esto deberá seguir este procedimiento:

- Ingresar a la página web del SiB (www.siac.net.co/sib) y dar clic en el *ícono archivador* en el menú de la izquierda, que se llama “Información de referencia”.
- En la página que se despliega a continuación (Herramientas de administración de datos), seleccionar el conjunto de referencia a utilizar en el menú desplegable: AAT.
- Digitar en el cuadro de búsqueda el nombre a consultar, o algunas de sus letras.
- Dar clic en “Buscar”. Como resultado de la búsqueda aparecerá, si está disponible, la ficha nomenclatural del nombre consultado (Figura VIII-1).

También es posible consultar, a través del “Catálogo de organismos y ecosistemas presentes en Colombia”, un nombre común asociado a un nombre científico y así, mirar su ficha nomenclatural. Para ello es necesario:

- Ingresar a la página Web del SiB (www.siac.net.co/sib) y dar clic en el vínculo “Catálogo de organismos presentes en Colombia”, que se encuentra en el área inferior derecha de la pantalla;
- en la página que se abre a continuación, digitar en el cuadro de búsqueda el nombre científico o común del taxón requerido (Figura VIII-2);
- seleccionar un taxón a consultar de las respuestas posibles y fichas disponibles.

En la página con los resultados de esta consulta aparecerá, si está disponible, una ficha de especie, donde en la primera línea aparece la información nomenclatural del taxón. Para consultar el detalle de la ficha nomenclatural, se debe dar clic en el triángulo invertido de “Información básica de la ficha” y nuevamente clic en el vínculo de “Ficha nomenclatural”.

¹³ <http://www.siac.net.co/sib/administracion2/WebModuleAdministracion/conjuntosreferencia.jsp>

¹⁴ <http://www.siac.net.co/sib/aat/WebModuleAAT/index.jsp>

Acidocroton

Autor: Griseb.
Reino: Plantae
Nivel taxonómico: Aceptado
Estado taxonómico: Aceptado
Taxones relacionados: Ophellantha (Relación sinonimia)
Estado de revisión/verificación: Revisado
Fecha de revisión/verificación: 2003-09-23
Comentario del especialista:
Distribución: No asignada

Jerarquía taxonómica:

Reino: Plantae Aceptado
División: Magnoliophyta Aceptado
Clase: Magnoliopsida Aceptado
Subclase: Rosidae Aceptado
Orden: Euphorbiales Aceptado
Familia: Euphorbiaceae Aceptado
Subfamilia: Crotonoideae Aceptado
Tribu: Codiaeae Aceptado
Género: Acidocroton Aceptado
Especie: Acidocroton gentryi Aceptado

Referencias

Fuente de la descripción :
Original: Grisebach, A. H. R. (1859) Flora of the British West Indian IslandsGoettingen 43pp
Secundaria: Webster, G. L. (1994) Synopsis of the genera and suprageneric taxa of EuphorbiaceaeAnnals of the Missouri Botanical Garden81 33-144

Fuente del estado taxonómico:
Original: Grisebach, A. H. R. (1859) Flora of the British West Indian IslandsGoettingen 43pp

Fuente del sistema de clasificación:
Original: Webster, G. L. (1994) Synopsis of the genera and suprageneric taxa of EuphorbiaceaeAnnals of the Missouri Botanical Garden81 33-144

Contacto:
 José C Murillo Aldana, jmurillo@ciencias.unal.edu.co, Cuidad Universitaria, Universidad Nacional de Colombia, Instituto de Ciencias Naturales

Figura VIII-1. Ficha nomenclatural del nombre “*Acidocroton*” cuando es consultado en línea en los AAT del SiB.

⇒ Descargas de AAT

Este procedimiento está diseñado especialmente para la descarga e integración de los AAT a las bases de datos estructuradas de acuerdo con el estándar de registros biológicos propuesto por el equipo coordinador del SiB. Sin embargo si no se cumple con este requisito se pueden descargar los conjuntos en *formato de texto* (.txt) e integrarlos a diferentes estructuras de bases de datos. Para descargar un AAT se deberá tener en cuenta lo siguiente:

- Solamente pueden descargar AAT los usuarios registrados en el sistema. Si no se ha registrado, hágalo dando clic en el *ícono lápiz* del menú de la esquina superior derecha y solicite autorización para descargar AAT al correo sib@humboldt.org.co.
- Ingrese a la sección “AAT disponibles” de la página web del SiB (<http://www.siac.net.co/sib/aat/WebModuleAAT/MetadatosAAT.jsp>) para saber si está disponible el AAT del grupo que se requiere descargar (Figura VIII-3).
- Dé clic en el título del AAT de interés para desplegar el metadato correspondiente.
- Dé clic en el *ícono disquete* al final del metadato (“Descargar”).
- En la ventana de respuesta, dé clic en el vínculo “login” y seguidamente proporcione su correo electrónico y su contraseña.

Una vez validada su información podrá descargar en el equipo local el AAT seleccionado.

⇒ Verificación de los nombres

Como se mencionó anteriormente, los archivos de autoridad taxonómica controlan el ingreso de errores nomenclaturales y facilitan la documentación de los registros biológicos en las bases de datos, puesto que los nombres científicos que se asociarán a los registros biológicos se seleccionan directamente del diccionario taxonómico. En los casos en los que no se cuente con esta herramienta, se puede incurrir en un gran número de inconsistencias nomenclaturales como el ingreso de nombres mal escritos (errores de tecleo) o de sinonimias. Por lo tanto se recomienda verificar cada uno de los nombres antes de ingresarlos en las bases de datos, por medio de diccionarios taxonómicos disponibles en línea. Adicionalmente a los que el SiB ha puesto a disposición, puede encontrar catálogos taxonómicos recomendados, organizados por grupo, en la sección de “Enlaces” del Sistema (http://www.siac.net.co/BancoConocimiento/S/sib_enlaces/sib_enlaces.php#AAT).

Entre los diccionarios recomendados están los del sistema “Biótica” de la Conabio (http://www.conabio.gob.mx/informacion/catalogo_autoridades/doctos/electronicas.html) y los de *Catalogue of Life* (www.catalogueoflife.org), que reúne los contenidos de *Species 2000* (<http://www.sp2000.org/>) e ITIS (<http://www.itis.usda.gov/>), que actualmente constituyen los listados de referencia más completos, actualizados y consultados de todos los disponibles en Internet. En el caso de la CONABIO, se pueden descargar diferentes diccionarios o catálogos al equipo local en formato de documento portátil (.pdf), hoja de cálculo (.xls) o archivo de base de datos de MS Access®. Basta con dar clic en el vínculo del formato seleccionado y el archivo se descargará a la ubicación que se indique. Para aprovechar mejor estos contenidos se recomienda leer detalladamente el modelo de datos de Biótica que corresponde al capítulo 12 del manual del programa (versión 4.1), en http://www.conabio.gob.mx/informacion/biotica_espanol/doctos/manual_v4.1.html.

En el caso de ITIS, los diccionarios se pueden consultar de múltiples formas. Primero se pueden hacer consultas sobre un taxón particular, con lo cual se obtendrá un reporte detallado que incluye datos sobre la nomenclatura y taxonomía, jerarquías taxonómicas, referencias, e información geográfica. Un segundo tipo de consultas incluye un reporte jerárquico y descargas personalizadas en diferentes formatos (véase: <http://www.itis.usda.gov/access.html>). Finalmente, en *Catalogue of Life* se pueden hacer consultas distribuidas de manera interactiva y dinámica desde <http://www.catalogueoflife.org/dynamic-checklist/search.php>. Cada nombre proviene de una fuente claramente identificada, a la que se puede acceder en línea dando clic en el vínculo correspondiente. Anualmente se generan también listados de referencia que se publican en CD-ROM.

⇒ Descripción de un conjunto de datos mediante el tesoro

En el Sistema de Información sobre Biodiversidad, el tesoro sobre biodiversidad permite describir conjuntos de datos biológicos y buscar información al interior del Sistema. Para ello debe seguir estos pasos:

- Ingrese a la página web del SiB (www.siac.net.co/sib) y dé clic en el *ícono archivador* en el menú de la izquierda, que se llama “Información de referencia”.
- En la página que se despliega a continuación (Herramientas de administración de datos), seleccione el conjunto de referencia a utilizar en el menú desplegable: Tesoro sobre biodiversidad.
- Digite en el cuadro de búsqueda el nombre a consultar, o algunas de sus letras.
- Dé clic en “Buscar”.

Catálogo electrónico de los organismos presentes en Colombia

Desplegar el árbol por:
 Nombre común Nombre científico

Plantas
 Animales
 Organismos

Buscar en el nombre científico o en el nombre común
 Buscar en el contenido de la ficha

Buscar

Búsqueda con filtro geográfico

Administración de contenidos

Entidades que aportan información:

Encuentre en el catálogo fichas de especies.
 Cómo buscar un organismo:
 Escriba en la casilla de búsquedas el nombre del organismo o del grupo biológico que desea encontrar y dé clic en **Buscar**.

También puede desplegar las ramas del árbol de nombres comunes o del árbol taxonómico dando clic en las carpetas. Cuando encuentre el organismo o el grupo de interés, dé clic en el nombre para transferirlo a la casilla de búsquedas y luego dé clic en **Buscar**.

Colombia Instituto de Investigación de Recursos Biológicos Alexander von Humboldt
 Equipo Coordinador del Sistema de Información sobre Biodiversidad y
 Proyecto Atlas de la Biodiversidad de Colombia - Programa de Inventarios de la Biodiversidad 2005. © Algunos derechos reservados. [Contáctenos](#)

Con el apoyo financiero de:

Powered By **ingenian** Software

Internet 100%

Figura VIII-2. Entrada del catálogo electrónico de los organismos presentes en Colombia, mostrando las posibilidades de búsqueda de nombres científicos.

Como resultado de la búsqueda aparecerá, si existe, el término buscado (o los términos que contengan las letras que se digitaron) en todos los contextos en los que ha sido catalogado dentro del tesoro (Figura VIII-4). Para seleccionar qué términos usar para sus descripciones, tenga en cuenta que no se encuentren precedidos de UP en el tesoro.

⇒ Búsquedas en el Sistema mediante el tesoro

Para buscar información al interior del sistema utilizando el tesoro haga lo siguiente:

- Acceda al tesoro sobre biodiversidad disponible en la página del SiB, en el URL: <http://www.siac.net.co/sib/tesauros2/WebModuleTesauros/index.jsp>.
- Ingrese el término (descriptor) de búsqueda.
- Seleccione los términos relacionados o equivalentes del tesoro y digítelos en el momento de buscar la información en cualquiera de los buscadores del Sistema (→ *Productos y servicios*).

Anteriormente era posible realizar búsquedas directamente a partir del buscador avanzado, usando los términos (descriptores) sugeridos en el tesoro. No obstante, esta función fue deshabilitada del portal SiB porque a través del Buscador súper-simple es más fácil acceder a todos los recursos. Se espera que

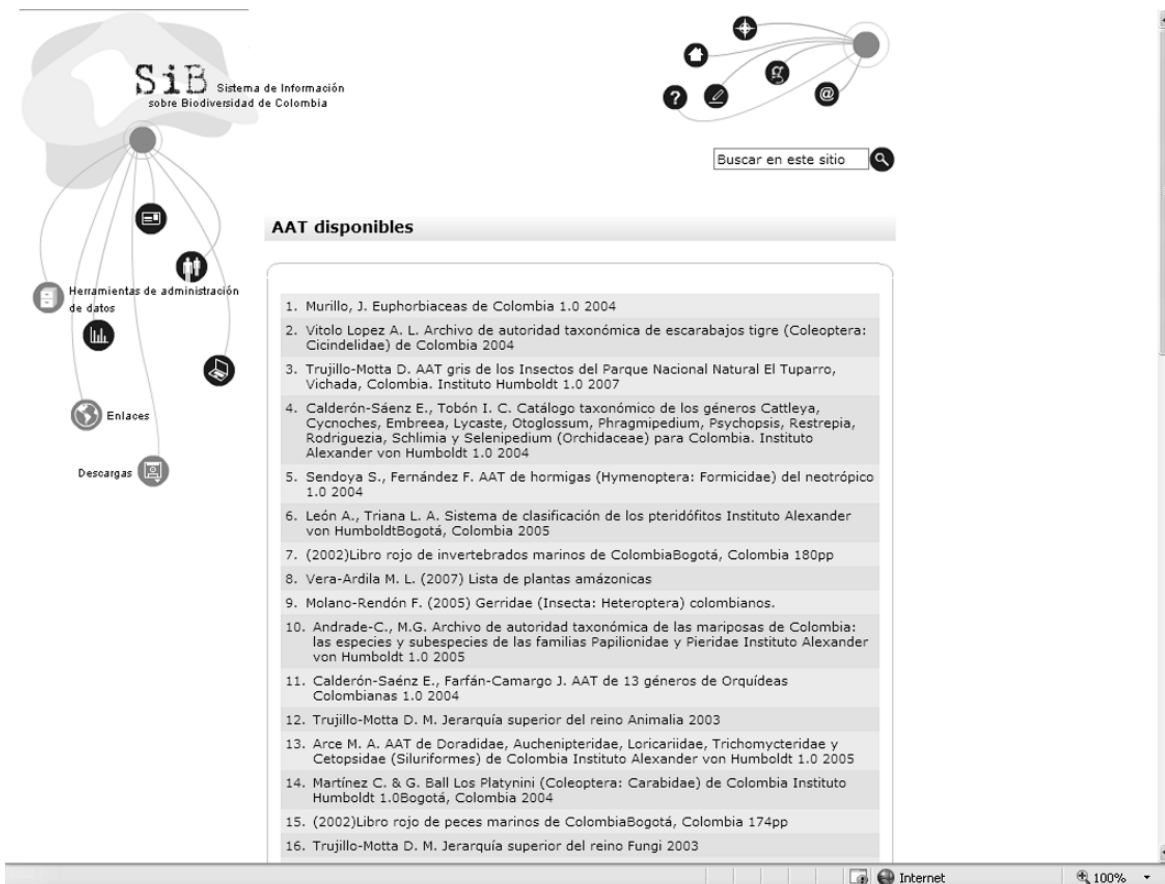


Figura VIII-3. Imagen de la pantalla de AAT disponibles.

próximamente, y en alianza con la *Red de información de especies y especímenes de IABIN*¹⁵ pueda proveerle a los usuarios del portal SiB la funcionalidad de búsquedas mediante tesauros, de manera que no solamente se busque por los términos elegidos sino por los equivalentes o relacionados que el sistema sugiere.

⇒ Consultas al catálogo de métodos y atributos


En la estructura actual del catálogo solamente es posible realizar búsquedas por palabras clave, suponiendo que estas consultas se hacen para responder preguntas específicas sobre las que se necesita ubicar toda la información disponible. Es recomendable entonces que las palabras o términos de búsqueda se hayan definido previamente a través del tesauro (→ *Búsquedas en el Sistema mediante el tesauro*). Para consultar el catálogo:

- Ingrese a la página web del SiB (www.siac.net.co/sib) y dé clic en el *ícono archivador* en el menú de la izquierda, que se llama “Información de referencia”.

¹⁵ IABIN es la Red Interamericana de Información sobre Biodiversidad y la Red de especies y especímenes (REyE en español o SSTN en inglés) es una de sus redes temáticas. La implementación de esta red se inició en 2005 y es coordinada por el Instituto Nacional de Biodiversidad (INBio) de Costa Rica. El Instituto Humboldt, a través del Equipo Coordinador del Sistema de Información sobre Biodiversidad, es socio del consorcio que lidera esta red.

Resultados de búsqueda en tesauros

7 registros encontrados por jagua Realizar nueva búsqueda [aquí](#)

Dé clic en  para ver términos específicos

MT (Microtesauro)	TG (Término general)	TE (Término específico)	TR (Término relacionado)	UP (Use por)
-------------------	----------------------	-------------------------	--------------------------	--------------

Este árbol de términos se encuentra en construcción, sus aportes o comentarios son bienvenidos y serán tenidos en cuenta para el desarrollo. Por favor no dude en escribirnos a: sib@humboldt.org.co












-  MT [Tesauro sobre biodiversidad](#)
 -  MT [Gacetero](#)
 -  TG [División político-administrativa](#)
 -  TE [Cesar](#)
 - TE [La Jagua de Ibirico](#)
 -  TE [La Guaiira](#)
 - TE [La Jagua del Pilar](#)
 -  TE [Sucre](#)
 - TE [Majagual](#)
 -  TG [Plantas](#)
 -  TE [Plantas vasculares](#)
 - up: [Plantas con flores](#)
 -  TE [Angiospermas](#)
 -  TE [Dicotiledóneas](#)
 -  TE [Rubiáceas](#)
 - TE [Jagua](#)

Figura VIII-4. Imagen de una consulta por el término “jagua” en el Tesauro sobre biodiversidad.

- En la página que se despliega a continuación (Herramientas de administración de datos), seleccione el conjunto de referencia a utilizar en el menú desplegable: Catálogo de métodos y atributos.
- Digite en el cuadro de búsqueda el término (método, atributo, tema) a consultar, o algunas de sus letras.
- Dé clic en “Buscar”.
- Seleccione la pestaña correspondiente a métodos o atributos.
- Elija los métodos o atributos a consultar.
- Expanda el término haciendo clic sobre el título (Figura VIII-5).

Nota: para encontrar las relaciones entre métodos y atributos, es necesario desplegar el término. El detalle de los atributos puede verse haciendo clic en cada uno de ellos desde la pestaña de su método relacionado o seleccionando la pestaña respectiva y dando clic en el título del atributo.

⇒ Documentación de métodos y atributos

La documentación de métodos y atributos es responsabilidad primaria de cada administrador de los conjuntos de datos (de acuerdo con su grado de participación), con la asesoría de los administradores de los conjuntos de referencia, concretamente del administrador líder. Sin embargo, debe recordarse que

The screenshot shows the SiB website interface. At the top left is the SiB logo and name. A navigation menu includes 'Metadatos', 'Archivos de autoridad taxonómica', 'Tesauros', and 'Métodos y atributos'. A search bar contains the text 'Buscar en este sitio'. Below the search bar, the page title is 'Resultados de búsqueda en métodos y atributos'. There are tabs for 'Métodos' and 'Atributos'. The main content area shows '26 métodos encontrados por aves'. The first method is '1. Puntos de conteo', which describes a field survey method involving two points, 50m radius, and 15-minute observations. The second method is '2. Captura mediante trampas Malaise', which describes a trap for flying insects using a mesh and preservative.

Figura VIII-5. Imagen de una consulta por el término “aves” en el Catálogo de métodos y atributos.

esta documentación es uno de los primeros pasos en el proceso mismo de investigación (→ ¿Cómo se obtiene la información?), por lo cual es muy recomendable que sea realizado independientemente de la afinidad del administrador con el Sistema.

Para incluir métodos y atributos en el catálogo del SiB, es necesario hacer un proceso personalizado. A diferencia de otros procedimientos, éste requiere del trabajo conjunto del administrador del conjunto de datos con el administrador del catálogo de métodos y atributos, debiendo el administrador cumplir los siguientes pasos (tomado de Suárez-Mayorga 2003):

- Elaborar un documento explicativo sobre el conjunto de datos que se está documentando, donde se identifique de la manera en que quiere ser conocido, se explique el propósito de su establecimiento y se definan claramente las metodologías empleadas para obtener los datos e información asociados al mismo, así como las entidades y atributos que lo conforman.
- Complementar la estructura de entidades y atributos diseñada por el administrador del Catálogo con base en el documento elaborado y las secciones 5 y 6 del Estándar para la documentación de registros biológicos (Suárez-Mayorga *et al.* 2005a).
- Discutir y definir conjuntamente la estructura de métodos y atributos propuesta y elaborar un documento final, que será aprobado por las partes.

Nota: Al obtener la aprobación final del administrador y tras el proceso de verificación de calidad, los métodos y atributos definidos serán ingresados al catálogo correspondiente y posteriormente, publicados en Internet.

ii. Documentación: datos multipropósito

a. Catálogo de metadatos

Los metadatos son un recurso de información que permite tener una descripción de los diferentes conjuntos de datos que están presentes en el sistema. Facilitan tanto la gestión como la búsqueda e integración de los contenidos de información. Uno de los puntos más importantes de los metadatos es que logran mantener los contenidos de información a lo largo del tiempo y los acontecimientos. De acuerdo con Michener *et al.* (1997), cada vez que se realiza una investigación, los contenidos de información van decayendo en relación con el tiempo de publicación (Figura VIII-6). El artículo de Michener y colaboradores propone que los metadatos pueden evitar la pérdida de información porque las descripciones considerarían aquellos detalles que son omitidos por los investigadores en el momento de realizar la publicación, manteniendo el contenido de información a pesar de que los datos cambien de administrador, o incluso muera el autor de la investigación o el generador del conjunto de datos.

Los metadatos desarrollan un papel protagónico también en el mantenimiento de la información a largo plazo. La preservación de la información siempre ha sido una preocupación de los responsables de ésta, porque tienen el reto de mantenerla a pesar de los cambios tecnológicos, tanto en software como en medios de almacenamiento. Se han propuesto diferentes alternativas para esto, sin embargo, todas involucran la documentación con metadatos, porque son el único mecanismo que asegura tener una descripción comprensiva tanto de los contenidos de información como de los medios y recursos que se utilizaron en la captura, el procesamiento y el almacenamiento de los datos.

En términos generales, puede decirse que los metadatos permiten:

- ~ Contar con un inventario de los conjuntos de datos generados y/o adquiridos.
- ~ Tener un mecanismo para mantener, controlar y auditar los conjuntos de datos, los usuarios y las consultas de información.
- ~ Conocer los contenidos y demás aspectos de los conjuntos de datos pertenecientes a otras organizaciones y grupos de trabajo.
- ~ Mejorar o actualizar los conjuntos de datos creados anteriormente, disminuyendo los costos de captura de información.
- ~ Evitar la repetición de datos previamente compilados por otros grupos u organizaciones.

Así mismo, los metadatos cumplen varias funciones importantes en el sistema:

- ~ Facilitan la comunicación a todos los niveles y entre todos los actores del Sistema de información.
- ~ Documentan el contexto en el cual los datos fueron capturados, de modo que facilita su reutilización en el futuro.
- ~ Reducen el riesgo de la devaluación de los conjuntos de datos por la pérdida de conocimiento sobre estos.

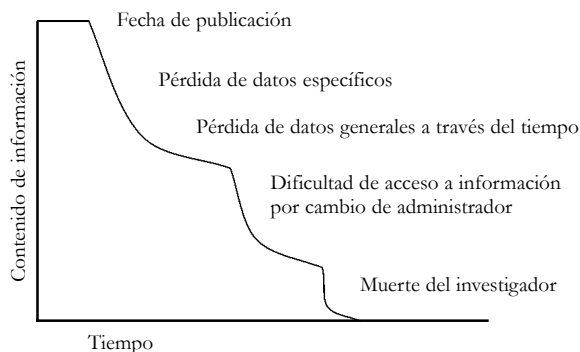


Figura VIII-6. Representación de Michener *et al.* (1997).

- ~ Facilitan la integración de conjuntos de datos de diversos grupos y organizaciones.
- ~ Facilitan las búsquedas de conjuntos de datos por tema, palabras clave, lugares geográficos o taxonomía.
- ~ Brindan una herramienta de acceso rápido para el descubrimiento y adquisición de información.

La captura de información en el catálogo de metadatos se realiza a través del proceso de documentación. La definición más sencilla de **documentar** es sustentar, probar o justificar mediante documentos; en el caso de los metadatos, se considera como el proceso de construcción de una descripción comprensiva de los conjuntos de datos disponibles en la organización. De acuerdo con el Federal Geographic Data Committee (FGDC 2000), los metadatos describen el contenido, la calidad, la condición y otras características de los conjuntos de datos en un formato estandarizado. La documentación es pues, el mecanismo a través del cual se alimenta el catálogo.

b. Procedimientos asociados

■ Documentación de metadatos

El procedimiento para documentar los metadatos es muy sencillo, desde que se cuente con los insumos necesarios para llevarlo a cabo. Realice los pasos marcados con guión:

- Identifique el conjunto de datos que va a documentar

La documentación se realiza sobre los conjuntos de datos o recursos de información (bases de datos, aplicaciones, documentos) que incluyan o estén relacionados con biodiversidad. Un **conjunto de datos** contiene la información que describe un grupo de observaciones colectadas mediante la misma metodología y cada conjunto es una unidad independiente de administración de información.

Para poder identificar adecuadamente cada conjunto de datos considere:

- ~ Las actividades realizadas para capturar la información.
- ~ La estructura de los archivos para guardar la información.
- ~ La metodología empleada para recoger la información.

Cabe recordar que los *conjuntos de datos* son también grupos de registros que tienen una coherencia lógica entre sí. Revise sus datos, vea cómo están agrupados y decida cuáles son en realidad conjuntos de datos, y cuáles no. Cuando tome esta decisión, organice la información en archivos separados, colocando nombre a cada conjunto identificado.

- Lea el estándar para documentar metadatos (en el caso del SiB el estándar vigente es el *Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad* (Rivera-Gutiérrez y Suárez Mayorga, 2005).

El estándar define la información que se debe considerar para documentar conjuntos de datos, indicando el elemento, el tipo de datos que se pueden ingresar y los posibles valores que puede tomar. Una adecuada lectura del estándar le permitirá comprender la información que necesita para documentar adecuadamente los metadatos de sus conjuntos de datos.

- Establezca los elementos que servirán de referencia para llenar la información

A partir del conocimiento del estándar, cada administrador puede establecer la información necesaria para realizar la documentación, a través de preguntas sencillas tales como: ¿qué información contiene el conjunto?, ¿para qué lo hago?, ¿quién ha intervenido en el conjunto de datos?, ¿cómo se construyó?, ¿cuándo?, ¿dónde? Se puede considerar la Tabla VIII-2 como una buena referencia.

- Lea la guía para usar la herramienta informática

En la página Web del SiB, sección “Descargas” (*ícono disquete* del menú izquierdo), carpeta “Metadatos”, subcarpeta “Documentos” encontrará una guía que le ayudará a utilizar la herramienta desarrollada por el EC-SiB para documentar metadatos: *Cassia*®. Es un texto sencillo que explica paso a paso como ingresar información en la herramienta para metadatos.

- Acceda a la herramienta informática

Cassia® está disponible en línea a través del portal SiB (www.siac.net.co/metadatos) o puede ser solicitada e instalada localmente, siempre que se firme una carta de compromiso en la que se estipula el uso que puede dársele (contacte para esto al Administrador del Sistema en sib_colombia@humboldt.org.co). Antes de ingresar a la herramienta debe registrarse como usuario (puede dar clic en el *ícono lápiz* que aparece en la página de documentación). El administrador de metadatos recibirá su solicitud y, después de confirmar sus datos, le indicará el procedimiento que debe seguir y le asignará un usuario y clave para que pueda hacer uso de la herramienta.

- Ingrese la información

Es de suponer que en este punto del proceso el administrador de información ya ha identificado sus conjuntos de datos, ha leído los diferentes documentos relacionados con la documentación de metadatos (estándar de metadatos, guía de usuario de la herramienta de documentación), ha preparado la información que servirá como base para la documentación y se encuentra con su computador encendido y con la herramienta lista para ingresar sus datos. Entonces, ya tiene los elementos necesarios para realizar la documentación, así que, ¡adelante!

Tabla VIII-2. Campos a documentar en metadatos dentro del SiB para Colombia.

¿Qué?	i. Título ii. Resumen iii. Palabras clave o descriptores iv. Taxonomía utilizada (sistema de clasificación, palabras clave) v. Entidades vi. Atributos
¿Quién?	i. Autor ii. Persona responsable de la información iii. Entidad responsable de la información iv. Créditos
¿Cómo?	i. Métodos utilizados ii. Calidad de los datos iii. Distribución a usuarios iv. Condiciones de acceso de los conjuntos de datos
¿Cuándo?	i. Tiempo asociado al conjunto
¿Dónde?	i. Referente geográfico que se corresponde con el conjunto de datos ii. Palabras clave geográficas iii. Coordenadas geográficas
¿Para qué?	i. Propósito ii. Público objetivo iii. Limitaciones y alcances de la información

Una vez hecha la documentación (y enviada, si no se hizo en línea), el administrador del “Catálogo nacional de metadatos” revisará la información y verificará que se hayan ingresado los datos mínimos solicitados y que lo que se ingresó en cada campo sea adecuado de acuerdo con el estándar. Si el metadato cumple con los parámetros de la verificación, será aprobado y se publicará automáticamente en el catálogo. Por el contrario, si no cumple con alguno de los criterios, se notificará al administrador responsable de la documentación para que revise y ajuste la información.

Para ilustrar mejor el procedimiento, a continuación se presenta un ejemplo que involucre todos los pasos del proceso. Suponga que usted (administrador) lideró una salida de campo a la Reserva Forestal de Yotoco en el Valle del Cauca. A la salida, además de usted, asistieron Felipe Arias, Catalina Rodríguez, Pedro Ruiz, Juan Carlos Silva y Gloria Díaz. La salida se hizo entre el 15 y el 30 de abril de 2003. Se aplicaron dos metodologías de muestreo, en dos transectos diferentes: una captura de aves utilizando 200 metros de redes de niebla y la otra, observación de aves con puntos. Las redes de niebla se abrían entre las 6 de la mañana y las 6 de la tarde y las observaciones se hacían en dos jornadas, primero en la mañana de 6 a 10 y luego en la tarde, entre las 3 y las 7 de la noche. Para el método de captura con redes se tomaron datos de morfometría en todos los ejemplares capturados, además del sexo, edad, estado del plumaje y coloración de partes blandas. Para las observaciones se registró el nombre de la especie, sexo, edad, actividad (alimentación, reposo, reproducción, vuelo), alimentación (insectos, néctar, semillas, frutas), estrato (arbóreo superior, arbóreo inferior, arbustivo, suelo, herbáceo) y sustrato (suelo, follaje, tronco y ramas grandes, ramas medianas, ramas pequeñas, aire, agua, sobre animales). Usted, como líder de la actividad, tiene los archivos en Excel© donde están organizados todos los datos.

Ahora es importante que identifique los conjuntos de datos que describen su actividad. Lo primero es que toda su salida de campo es un gran conjunto de datos, el cual debe ser descrito de manera indepen-

diente, y será el conjunto de datos “padre”, en el cual se pueden circunscribir otros. El conjunto de datos padre define el evento de captura de información, en el cual se establecen los colectores, el lugar de trabajo, el tiempo y los métodos aplicados. Ahora, ¿cómo identificar los conjuntos de datos subordinados o “hijos”? La definición de los mismos puede realizarse por los métodos de captura que se utilizaron; en este caso, captura con redes de niebla y observaciones por transecto. Así, tendríamos tres conjuntos de datos identificados, uno padre y dos hijos (Figura VIII-7).

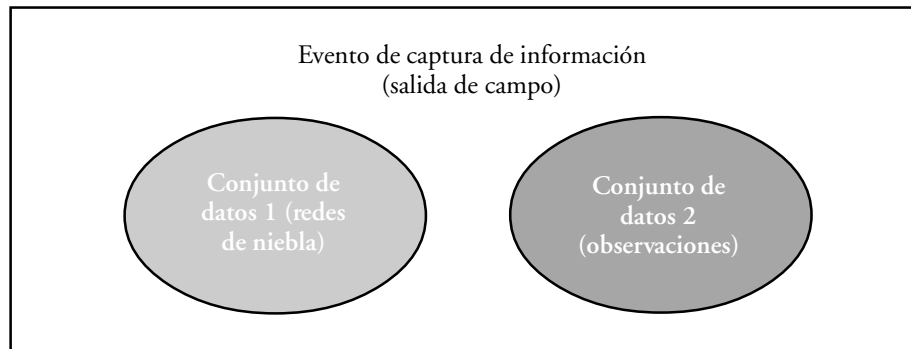


Figura VIII-7. Ejemplo práctico de definición de conjuntos de datos.

IX. ¿Cómo se almacena la información bajo los parámetros SiB?

Nelson Andrés Ramón-Rodríguez, Ángela M. Suárez-Mayorga y Ximena Franco-Villegas

i. Registros biológicos

Partiendo de que en el Sistema la mayoría de datos e información hacen referencia a unidades biológicas, y con base en ello se han construido herramientas conceptuales como los estándares de registros biológicos, no es sorprendente que dichos datos e información se almacenen bajo una estructura denominada “registros biológicos”. A continuación se explica detalladamente la conceptualización de esta estructura.

Para comenzar, es necesario definir ¿qué es un registro biológico? Un **registro biológico (RRBB)** es la observación y captura de información sobre **unidades biológicas**, referidas a un espacio y, por lo general, un momento o lapso determinado. Uno de los conjuntos de registros biológicos más completo lo constituyen las colecciones biológicas. En la medida en que para el **registro biológico** la unidad biológica se refiere a un espacio y tiempo determinado y se caracteriza por unos atributos determinados, los registros biológicos están compuestos de:

- ~ **referente conceptual:** identificación y descripción mediante atributos de la unidad biológica,
- ~ **referente espacial:** referencia al espacio descrito para la unidad o unidades registradas;
- ~ **referente temporal:** referencia al momento de un proceso o registro, y no existe en todos los casos. Como ejemplo se pueden mencionar las fechas y horas de un muestreo de aves mediante redes de niebla (→ *Catálogo de metadatos: Procedimientos asociados*), donde importa el tiempo de extensión de la red, o la fecha de colección puntual de un ave que se preserva para museo;
- ~ **fuentes:** la referencia al responsable del registro total o de sus partes (p. ej. valores de los atributos, identificaciones, colección de la evidencia física, entre otros).

A su vez, el **referente conceptual** incluye:

- ~ **Identificación de la entidad:** cómo se designa la unidad y en qué jerarquía o sistema de clasificación se encuentra. También se conoce como **referente de clasificación**.
- ~ **Identificación y descripción de la evidencia física:** cómo se reconoce la unidad biológica a partir de evidencias tangibles de su registro.
- ~ **Definición de métodos y atributos:** detalle de las características particulares de la unidad o unidades registradas.

En la Figura IX-1 se ilustran los contenidos de cada referente.

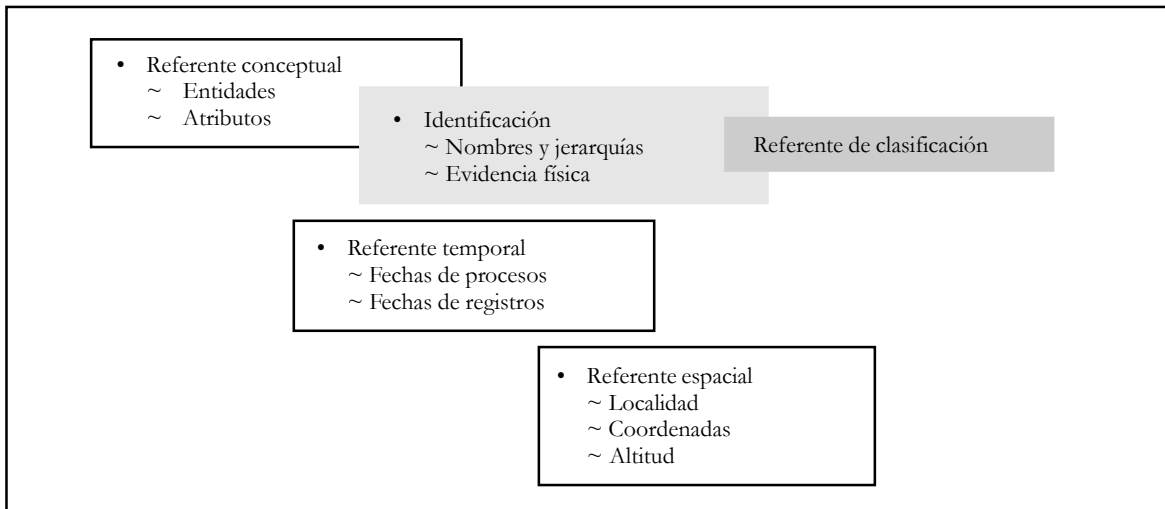


Figura IX-1. Descripción de los componentes de un registro biológico.

Para terminar, las fuentes pueden ser citas (referencias bibliográficas) o contactos (personas u organizaciones), que en el estándar son secciones recurrentes (→ *Estándares*). Ejemplos de contactos son los colectores de los especímenes, o las colecciones donde se guardan, mientras que ejemplos de citas son los documentos que soportan una metodología.

ii. Estructuras recomendadas de base de datos

En la presente sección se explicará la estructura de base de datos relacional que se propone para los registros biológicos (RRBB), sus diferentes módulos y posibilidades de uso para desarrollar software que la administre de forma personalizada. La estructura fue desarrollada a partir del Estándar para la documentación de registros biológicos¹⁶, y siguiendo el enfoque del Estándar para documentación de metadatos de conjuntos de datos relacionados con biodiversidad¹⁷, la Propuesta para la implementación de los tesauros sobre biodiversidad en el IAvH¹⁸, el Estándar para la elaboración de archivos de autoridad taxonómica¹⁹ y el Documento explicativo del catálogo de métodos y atributos sobre biodiversidad²⁰.

Como se observó en el apartado anterior, la documentación de un registro biológico se basa en 3 referentes y dos componentes de referencia, que surgen de las siguientes preguntas: ¿donde se registró?, ¿cuando se registro?, ¿cómo se registró?, ¿qué se registro? y ¿quién lo registro? El estándar para la documentación de RRBB expone, de una forma estructurada, los elementos que tienen estos 5 referentes, cuales de éstos

¹⁶ Suárez-Mayorga A. M., Rivera-Gutiérrez H. F., Varón-Londoño A., Ramón N. A. 2005. Estándar para la documentación de registros biológicos, versión 5.0. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D. C., Colombia, 64 p.

¹⁷ Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad; Héctor Fabio Rivera Gutiérrez, Ángela M. Suárez Mayorga, Andrés Varón Londoño. Sistema de Información sobre Biodiversidad, Instituto de Investigación de Recursos Biológicos "Alexander von Humboldt", revisado y aumentado en Rivera-Gutiérrez H. F. & Suárez-Mayorga A. M. 2005. Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad, versión 2.0. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D.C., Colombia, 66 p.

¹⁸ Propuesta para la implementación de los tesauros sobre biodiversidad en el IAvH; Talía Waldrón.

¹⁹ Trujillo-Motta D., Varón-Londoño A., Bello-Silva J.C. 2005. Estándar para la elaboración de archivos de autoridad taxonómica. Versión 5. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D.C., Colombia, 51 p.

²⁰ Documento explicativo del catálogo de métodos y atributos sobre biodiversidad; Ángela M. Suárez Mayorga.

son obligatorios para un RRBB y cuales son de valor siempre y cuando se tenga disponible, lo cual puede generar un esfuerzo para obtenerlos y registrarlos.

Estos referentes se ven reflejados en la estructura de la siguiente manera:

- ~ **¿Dónde se registró? o referente espacial:** se encuentra en las tablas que se vinculan con la tabla llamada “referente geográfico”, la cual incluye una ubicación político administrativa, topónimos socioculturales (fincas, barrios, reservas naturales) y topónimos físicos (cuchilla, monte, valle, laguna), de manera que la ubicación del registro corresponde al área de intersección entre estos elementos. También incluye la información de un sitio que puede estar relacionado con varias georreferencias (tomadas con métodos documentados) e instrucciones de acceso al mismo. Esto es equivalente a un punto o una serie de puntos donde fueron tomados los registros, adentro de un área ya delimitada. Cada georreferencia también tiene un componente altitudinal asociado (tomado con métodos documentados).
- ~ **¿Cuándo se registró? o referente temporal:** se encuentra en la estructura como dos elementos adentro de la tabla de RRBB, en la forma de intervalo de fechas (fecha inicial y final) e intervalo de horas.
- ~ **¿Cómo se registró? Referente conceptual - métodos:** se encuentra en la estructura en las tablas que manejan los métodos por los cuales se logra obtener un RRBB o se evalúan los atributos definidos para la entidad.
- ~ **¿Qué se registró? Referente conceptual - referente de clasificación y atributos:** se encuentra en la estructura en las tablas que manejan el referente de clasificación para un RRBB, así como en las tablas del componente de atributos, que a su vez diagnostican la identidad definida en el referente de clasificación. Cada unidad biológica puede tener uno o varios referentes de clasificación, que se asocian siempre a una fecha y a una fuente responsable.
- ~ **¿Quién lo registró? Referente fuente:** Esta información se registra en la tabla de metadatos (para el conjunto de datos) o directamente relacionado con un RRBB, bien en la tabla que almacena la información de los individuos que hayan participado en la colección, identificación o administración del registro biológico (contactos), o bien en la tabla que almacena la información de documentos fuente (citaciones).

En la Figura IX-2 se ve el diagrama conceptual de los registros biológicos. Las uniones punteadas encerrando los elementos y con los rectángulos como vértices, son las uniones entre los conjuntos de datos de referencia (→ *Conjuntos de referencia*). Los rectángulos son los conjuntos de datos de referencia que enmarcan la información de un RRBB. Los paralelogramos son los repositorios de datos recurrentes (recurrentes porque se recurre a ellos muchas veces, véase la explicación en → *Estándares*), y los círculos son los referentes de los registros biológicos, que están ligados a conjuntos de datos de referencia mediante uniones sólidas.

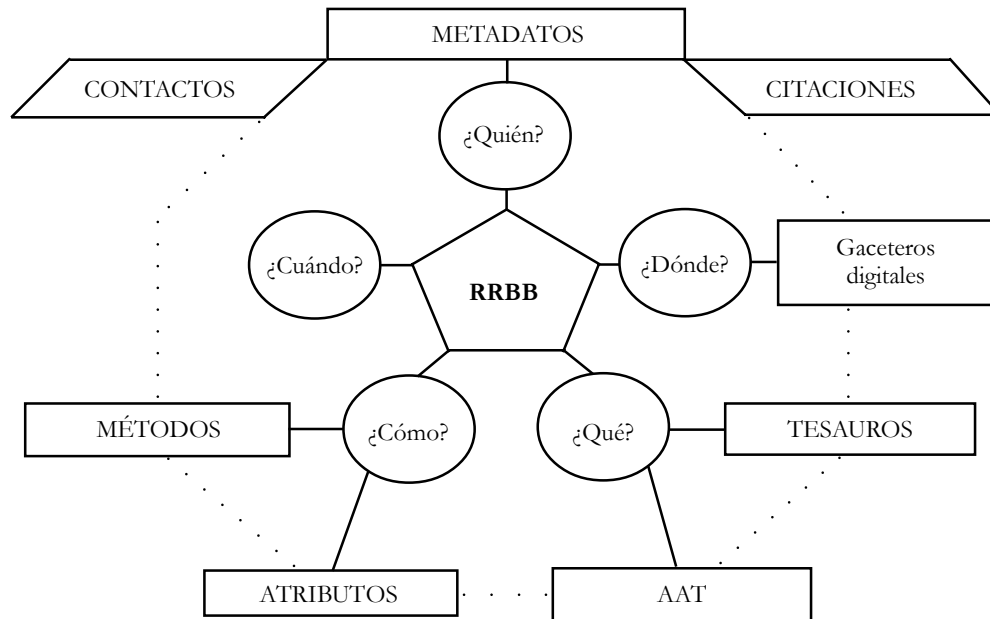


Figura IX-2. Diagrama conceptual de la estructura recomendada de base de datos para capturar registros biológicos.

La forma como fueron diseñados los referentes para ser representados en la estructura hace que su funcionalidad se pueda extender de acuerdo con las necesidades de las personas o grupos que la vayan a utilizar. La fortaleza de la estructura de base de datos relacional para registros biológicos desarrollada bajo el marco de los estándares del SiB radica en la *robustez* y el *dinamismo*.

Se dice que la estructura es **robusta** en cuanto a que soporta el manejo de datos e información de cualquier proceso de investigación o colecciones o información que tenga que ver con un registro biológico, y se dice que es **dinámica** ya que como tal, la estructura no posee un programa o *software* que se encargue de administrar la información que se ingresa, por lo que se puede hacer una herramienta personalizada. Otra fortaleza en el diseño dinámico de la estructura radica en que se puede personalizar la misma información que se tenga disponible adentro de la base de datos como insumo para el trabajo.

La estructura maneja además referentes básicos, información valiosa para un RRBB, como la que hace alusión a las evidencias físicas tomadas y sus repositorios, las citaciones pertinentes de los métodos y atributos para tomar el RRBB, teorías o sistemas de clasificación utilizados, la información del responsable (s) (definida como contacto: persona o entidad) de registrar, ubicar, clasificar, almacenar y administrar un conjunto de RRBB, y una especie de ficha bibliográfica del conjunto de datos, esta “ficha” o etiqueta con la descripción de los elementos del conjunto y del conjunto como tal, es el metadato del conjunto de datos (→ *Documentación: datos multipropósito*).

Otra fortaleza de la estructura propuesta está en el manejo que se le puede dar a los metadatos. Dada la importancia de los mismos (→ *Catálogo de metadatos*), todo registro biológico en el marco del SiB debe estar ligado a un metadato. Sobre los metadatos se enfocan las búsquedas y a través de ellos se hacen visibles tanto el conjunto de datos como su responsable o custodio y la manera de contactarlo, ya que, como se ve en el diagrama, los metadatos están a la cabeza de todo pero no pueden estar sin los conjuntos de referencia. En esta estructura, la versatilidad que se puede obtener con el concepto de metadatos tiene que ver con la “contenencia”, es decir, la posibilidad de incluir unos metadatos dentro de otros metadatos

y así crear jerarquías de metadatos. El metadato superior (metadato “papá” o más incluyente) se puede equiparar a una bolsa que contiene cajas, que a su vez contienen objetos específicos, que a su vez están conformados por piezas independientes (metadatos “hijos”).

Esta forma de establecer jerarquías permite que la estructuración o la concepción de la agrupación y contención de los conjuntos de datos pueda ser modelada según la necesidad o el gusto del investigador, según como plantee su investigación y como quiera trabajar sus datos, pero siempre dentro de los parámetros del Sistema. De esta forma se construye una especie de árbol n-ario (n hijos), donde sus nodos son un metadato con sus registros, los cuales pueden tener cuantos nodos hijos quieran. Sabiendo esto, se puede decir que de un metadato se puede derivar información basándose en los datos de sus registros, evitándose un paso de digitalización adicional; a la inversa, los datos de un metadato “papá” se pueden obtener a partir de sus metadatos hijos y de sus registros.

Se señalan ahora las ventajas ofrecidas por la estructura, analizándolas referente por referente:

- ~ **Referente espacial:** el diseño de las tablas que manejan este referente permite que se le pueda dar un manejo internacional (y no únicamente colombiano) a la estructura. Esto se logra gracias al uso de estándar ISO en cuanto a siglas de países y ciudades, que se generalizó a país, subdivisión administrativa (no únicamente “departamento” pues en otros países las divisiones administrativas tienen nombres diferentes como estado o provincia), ciudad-municipio y población (que tiene diferentes tipos de población como caserío o inspección de policía). Además de esto, para Colombia se incluye el código DANE para el manejo de las divisiones administrativas y los niveles en los que se encuentran. Con esta estructura es posible tener georreferenciados tanto la división política como los topónimos, para poder hacer análisis SIG de una forma más fácil. La georreferenciación se recomienda que sea almacenada en formato decimal y no hexadecimal, ya que el formato decimal es el que usan las aplicaciones SIG.

En conjunto con los metadatos se puede crear un esquema de trabajo y uso de la estructura (con *software* personalizado), de tal forma que los RRBB tengan por defecto un referente geográfico y sitios predefinidos y que no sea necesario ingresarlos nuevamente cada vez. Se puede derivar la información de las coordenadas de extremos norte, sur, este y oeste (que se requiere en el metadato), o desde las georreferencias de los RRBB hacia el metadato que los contenga.

- ~ **Referente temporal:** en conjunto con la tabla metadatos, se puede derivar información en dos direcciones según la necesidad y el diseño de un proyecto o forma de trabajo. Las opciones son: derivar la información desde los datos de cada RRBB hacia los metadatos para lograr tener el rango de fechas que los abarque a todos sin tener que digitarlo, u obtener fechas para los registros desde el metadato hacia sus RRBBs.
- ~ **Referente conceptual:** esta es una de las partes dinámicas de la estructura, ya que se pueden tener tantos métodos como se desee, cada uno con cuantos atributos se quiera y estos atributos se pueden definir con su propio dominio de valores, de tal forma que se adapten a cada necesidad. Los atributos pueden ser anidados (p. ej. el administrador puede escoger a,b,c para el valor 1 del atributo X y también puede escoger d, e, f para el valor 2 del atributo X) ya que, en esencia, todos son atributos que describen unidades biológicas y, por lo general, las unidades biológicas están anidadas entre sí. También existe una tabla para poder agrupar en subconjuntos

los métodos que una persona o grupo defina (p. ej. GEMA: insectos, aves, plantas...), de tal forma que el trabajo sea más organizado.

- ~ **Referente de clasificación:** esta parte también ofrece dinamismo. Se puede clasificar un RRBB con palabras clave que se buscan en un tesoro (como por ejemplo nombres comunes), se puede clasificar un RRBB con un nombre taxonómico proveniente de un AAT o se puede generar un sistema de clasificación diseñado a la medida de la necesidad. Es posible hacer clasificaciones que tengan una estructura jerárquica (p. ej. un árbol genealógico) a partir de la estructuración estandarizada de la información y se pueden categorizar estas jerarquías, lo cual permite, por ejemplo, decir qué familia incluye cuál género o qué género incluye cuáles especies. Existe la funcionalidad de asignar cuantos referentes de clasificación se quiera a un único registro biológico, siempre que estén asociados a una fecha y a una fuente (contacto o referencia bibliográfica) responsable.
- ~ **Contactos:** se puede llevar un registro de responsables por cada decisión que se tome en el marco de un RRBB, tal como el responsable de la identificación, el que tomó el dato, el custodio de una evidencia, el autor de un método, el autor de un atributo, la persona que tomó la georreferencia, etc. En este *software* estas relaciones se pueden manejar de forma personalizada, como por ejemplo, al asignar valores por defecto para todos los RRBB o georreferencias.

Entonces, las oportunidades que se ofrecen con el diseño de esta estructura de base de datos son las de poder desarrollar aplicaciones personalizadas para manejar la misma estructura de base de datos, aunque respondiendo a necesidades específicas; la clave es que todo se desarrolla a partir de una única estructura. A su vez, esto permite personalizar información adentro de la estructura, con el fin de que esté disponible para el trabajo de un usuario en especial (p. ej. una colección biológica, un herbario o un parque nacional natural).

iii. Seguridad y almacenamiento físico

Tomado de Franco Villegas 2003: "Política institucional para la gestión de datos e información sobre biodiversidad". Revisión: Andrés Ramón.

Respaldar la información consiste en duplicarla como medida preventiva, para poder acceder a la copia en caso de que el original se pierda o dañe. Estos daños pueden surgir por diferentes razones tales como:

- ~ Errores de aplicación (*software*): una aplicación puede fallar por diversas razones, bloqueándose o dejando de funcionar sin poder salvar el trabajo realizado desde el último procedimiento.
- ~ Fallos de *hardware*: un disco duro puede fallar de tal forma que se dañen sectores de memoria, que pueden ser recuperados en contadas ocasiones o cuando la caída de la cabeza de lectura deja completamente inservible el disco. Estos daños físicos en la memoria ocasionan por lo general la pérdida total de los datos grabados en ella.
- ~ Errores de usuarios: un usuario puede borrar archivos accidentalmente.
- ~ Errores del personal de administración del sistema: son comparables con un error de usuario, salvo que un administrador está a cargo de un volumen mucho mayor de datos y de mayor relevancia para la organización. Un ejemplo es la eliminación accidental de una cuenta en un servidor.

- ~ Accesos no autorizados: sucede cuando personas desconocidas para el sistema acceden al repositorio de información a través de una red o directamente en el puesto de trabajo, para alterar los datos.
- ~ Virus informáticos: pueden ser los mayores causantes de pérdida de información y daño de equipos, ya que usualmente atacan sin poder ser detectados; además se duplican y se autoenvían a otros equipos vecinos a través de una red (Internet, aplicaciones de *chat*) o por la transferencia de información por medios físicos de almacenamiento infectados (discos, CD's, memorias USB).
- ~ Accidentes: puede suceder, por ejemplo, que se riegue líquido en el equipo (es más común de lo que se creería), o que se presente una subida de voltaje intempestiva.
- ~ Robos y desastres naturales.

Estos eventos no siempre pueden ser previstos y evitados por un usuario. Si alguna de estas amenazas se materializa, habrá graves consecuencias que van desde pérdida parcial de datos hasta daños totales de datos y equipos.

Por lo tanto, es imperioso realizar copias de seguridad de los datos e información al cuidado de los administradores periódicamente, dependiendo de la frecuencia de actualización de sus conjuntos de datos.

El costo comparativo de crear una copia de los datos contra su pérdida total hace que el adoptar prácticas y políticas de seguridad de redes y equipos de trabajo, al igual que sistemas de copias de seguridad (*backup*) eficientes y confiables, sea barato y se justifique.

Tomar medidas de seguridad en redes y equipos (ya sea que estén en una red o aislados), generar cultura, implantar políticas e implementar sistemas de copias de seguridad es una práctica obligada dentro de la administración de seguridad de un sistema informático. Dentro de las medidas de seguridad que se pueden adoptar se encuentran:

- ~ Mantener antivirus con las actualizaciones pertinentes y más recientes.
- ~ Adquirir pólizas de seguros con gran cubrimiento para los equipos.
- ~ Instalar claves de acceso para los equipos personales, o permitir el bloqueo del equipo cuando el operario asignado no este presente.
- ~ Incorporar sistemas de copias de seguridad que incluyan varias copias en más de un sitio, de forma que los sitios estén espacialmente separados entre sí (al menos a 5 km de radio).

iv. Procedimientos asociados

Pensar en una copia de seguridad es garantizar un respaldo de la información (datos). La publicación de información y la gestión de los datos en un ambiente distribuido es de por sí una forma de garantizar duplicados de seguridad de la información, pero también es necesario establecer políticas para la generación de copias de seguridad físicas de los datos, pues representan un capital fundamental dentro de una organización.

a. Copia de seguridad para bases de datos

Para establecer una copia de seguridad se deben tener en consideración los siguientes pasos:

- Identificar la información relevante. La información personal NO ES relevante para la organización; únicamente se deben tener en cuenta la información y las aplicaciones vitales para el completo funcionamiento del sistema.
- Crear una carpeta nueva donde se realice siempre la copia de seguridad (“Copia de Seguridad”).
- Crear dentro de “Copia de seguridad” otra carpeta con la fecha en la que se hace la copia de seguridad y el nombre o acrónimo de quien está realizando la copia (por ejemplo, “20070828-ASM”).
- Copiar los archivos identificados en el primer paso a dicha carpeta.
- Realizar este proceso tan seguido como la actualización de la información lo amerite.
- Copiar esta carpeta en diferentes lugares físicos (en una carpeta de un compañero de trabajo, en la casa, en un servidor de *backups*, entre otros).
- Se recomienda que las copias de seguridad sean grabadas en medios magnéticos, rotuladas igual a las carpetas creadas anteriormente. La frecuencia de dicho copiado se hará de acuerdo con la regularidad de los cambios de información.

X. ¿Cómo fluye la información?

Julia Benavides Molineros, Ángela M. Suárez-Mayorga, Héctor Fabio Rivera-Gutiérrez y Ximena Franco-Villegas

i. Definición de condiciones de acceso y uso

La REMIB menciona dos tipos de acceso a los datos:

- ~ **Acceso libre:** cuando toda la información depositada en un conjunto de datos puede ser vista y utilizada por cualquier usuario; puede tratarse de datos ya publicados anteriormente, por lo cual los derechos sobre ellos están claramente definidos y comprobados.
- ~ **Acceso restringido:** cuando se decide no “liberar” parte de la información relacionada, por diferentes causas, como por ejemplo:
 - La divulgación de la información pone en peligro la conservación de la biodiversidad
 - Los datos se encuentran en proceso de análisis con miras a la publicación.
 - El *administrador* de los datos tomó su información de terceros que impusieron restricciones sobre el uso de la información.

En caso de que los datos no hayan sido publicados, el temor por problemas con los derechos de propiedad sobre los datos (plagio, por ejemplo), es real. Sin embargo, y dado que los datos tienen en ellos mismos un valor intrínseco para muchas personas y entidades además del investigador que los administra, deben establecerse reglas para incentivar la publicación de datos en un tiempo máximo a partir de su recolección y digitalización.

Para que un administrador defina las condiciones de acceso y uso de sus datos, debe tener en cuenta, entre otros, los siguientes criterios:

- El impacto que el acceso a sus datos o información pueda tener sobre las entidades que se describen en ellos (como por ejemplo, el impacto de conocer la distribución de especies amenazadas por comercialización ilegal).
- El impacto que el acceso a sus datos o información pueda tener sobre terceros.
- La publicación previa de los datos en otros medios, lo que daría lugar a un registro de autoría de los datos.
- Si los datos han sido obtenidos en el marco de convenios interinstitucionales, deben tenerse en cuenta los términos y condiciones establecidos en ellos (Tabla X-1).

Tabla X-1. Descripción de los tipos de acceso y uso de los datos e información

PROCESOS	CATEGORÍAS DE USO	
	LIBRE	RESTRINGIDO
Acceso	Libre	Condicionado por responsable, solicitud de autorización por diversas vías, inscripción previa, confirmación de datos
Uso	Libre uso, siempre y cuando cite adecuadamente la fuente	Condiciones definidas por la fuente y para los propósitos determinados por ella (P. ej. Condición: permiso escrito)

ii. Acuerdos y convenios

Para definir las condiciones de acceso y uso de la información es necesario tener claridad del marco de referencia de la captura y almacenamiento de los datos. Es decir, deben considerarse los contextos posibles para el acceso y uso de la información, como por ejemplo, el contexto político (acuerdos, convenios o contratos que pueden estar delimitando el uso y propiedad de la información), el contexto técnico (tipo de información y condiciones de la misma), el contexto económico (beneficios económicos que puede generar el acceso o uso de la información y distribución de beneficios), y el contexto social (necesidad y propósitos de la información).

En el marco del SiB es muy importante que el administrador defina para sus conjuntos de datos acuerdos o convenios que regulen el acceso, el uso o el intercambio de los mismos.

Semánticamente un **acuerdo** es sinónimo de un **convenio**, definidos como algo a lo que se accede conjuntamente. Sin embargo, por lo general los convenios representan una relación más estricta. BCIS (2000c) propone la realización de dos tipos de documento: acuerdos de puesta a disposición de datos y memorandos de entendimiento. Los **acuerdos de puesta a disposición** son documentos por los cuales tanto el administrador como el usuario de los datos e información manifiestan sus inquietudes y resuelven aceptar lo que la otra parte propone, teniendo como base la confianza entre las partes (que queda manifiesta en el documento). Una ventaja de estos acuerdos es que permiten establecer y hacer explícita la propiedad intelectual de los datos e información que se pongan a disposición. Un **memorando de entendimiento**, además de lo anterior, hace referencia a las disposiciones legales y económicas que rigen el acuerdo firmado entre las partes. En el capítulo 3 de BCIS (2000c), se presenta un formato de memorando de entendimiento para el acceso y uso de datos e información.

Para realizar un acuerdo de puesta a disposición de datos e información es útil considerar los siguientes parámetros (adaptado de BCIS 2000c):

- ¿Qué tipo de información se maneja? ¿Hay alguna que merezca ser restringida?
- ¿Qué cantidad de usuarios puede acceder y usar los datos?
- ¿La definición de condiciones de acceso y uso depende de alguien (persona o entidad) diferente al administrador? ¿Hay otros acuerdos o convenios vigentes sobre los datos e información que contiene el conjunto de datos a ser accedido?

- ¿Qué tipo de usuarios (investigadores, tomadores de decisión, entre otros) puede acceder a los datos y utilizarlos?
- ¿El uso de los datos generará algún beneficio económico? Si es así, ¿qué retribución tendrá el administrador por permitirlo?
- ¿Los datos e información pueden ser usados por terceros? ¿Qué tipo de terceros?
- ¿Qué derechos tendría el administrador sobre los productos de información fabricados a partir de los datos a su cargo? Por ejemplo, el administrador debe recibir siempre los créditos correspondientes al conjunto de datos utilizado y copia de todas las publicaciones que de él se deriven.
- ¿Cómo son los créditos del conjunto de datos? (→ Catálogo de metadatos).
- ¿Qué pasa si el usuario o el administrador incumplen el acuerdo? ¿Qué disposiciones legales aplican?

Consideremos los siguientes ejemplos: un investigador está contratado para realizar un estudio sobre el uso que hacen las comunidades de una población humana en la Guajira, de una especie de tortuga, y en el contrato dice que la entidad contratante será dueña de los derechos patrimoniales de la información. ¿Cuáles son las condiciones de acceso y uso? La respuesta es sencilla: son las que defina la entidad contratante. En el mismo ejemplo, pero suponiendo que la información se tomó en un convenio con una Corporación Autónoma Regional y que el acceso y uso de la información estará restringido, ¿cuáles son las condiciones? Bajo estas circunstancias, la decisión depende también de la aprobación de la Corporación para definir qué se hace con la información.

Ahora, supongamos que el administrador puede decidir sobre el uso de la información en su organización, pero ahora la respuesta al acceso y uso no está ligada a los acuerdos, convenios o contratos, sino a la pertinencia de publicar cierto tipo de información. Así que debe considerarse que información se está publicando, cómo afecta a los intereses de la entidad, de los investigadores o de las comunidades que explotan el recurso el que se haga pública la información, cuál es la relevancia de la información que se pone disponible, cuales con las implicaciones sociales, culturales, políticas y económicas alrededor de la publicación de la información, más aún, como podría verse afectada la unidad biológica. En el ejemplo, deben decidirse las condiciones de acceso y uso de un conjunto de datos que describe las localidades de especies de fauna que están amenazadas por comercio ilegal. ¿Qué puede suceder si la información es adquirida por comerciantes ilegales de fauna? En este caso la consideración normal es que las condiciones de acceso y uso sean restringidas. Igual puede suceder con información de uso y/o prácticas de manejo que realiza una comunidad particular sobre un recurso biológico, donde las condiciones de acceso y uso de la información dependen de la negociación realizada con la comunidad que está cediendo su información. No obstante, un argumento poderoso en contra de cualquier restricción para los conjuntos de datos puestos en Internet es que quienes hacen un uso inapropiado de la información disponible, por lo general pueden obtenerla sin la necesidad de que esté disponible en Internet; es decir, no necesitan de la publicación de la información “restringida” para hacer mal uso de ella.

En consecuencia, desde 2005 el Sistema de Información sobre Biodiversidad de Colombia adoptó para su portal Web una licencia Creative Commons (CC), que incluye algunos derechos reservados y que garantiza a los administradores de información, y a los usuarios del portal, un reconocimiento apropiado de las autorías y créditos de la información distribuida a través del portal. Así mismo, esta licencia advierte de usos indebidos (como para fines comerciales) y de la obligación de compartir, bajo la misma licencia, todos los productos derivados de la información disponible en el portal SiB. Para leer el detalle de esta licencia dé clic en el *ícono CC* de la sección inferior del *Home* del portal SiB (www.siac.net.co/sib) o consulte el vínculo <http://creativecommons.org/licenses/by-nc-sa/2.5/deed.es>.

iii. Productos y servicios de información

Los **productos** son recursos de información diseñados para una audiencia específica y con un propósito definido, y son resultado de la compilación y presentación de información analizada o interpretada (Franco Villegas 2003).

Los conceptos de productos y servicios no se pueden distinguir fácilmente. Para establecer una diferencia resulta útil revisar las definiciones de cada término que aplican para este contexto:

producto: cosa producida (RAE 2001).

servicio: función o prestación desempeñadas por organizaciones y su personal. Prestación humana que satisface alguna necesidad social y que no consiste en la producción de bienes materiales (RAE 2001).

Entonces, un producto se asocia con una cosa concreta y material, mientras que un servicio se asocia con una función. Para el caso de los productos y servicios de información, puede decirse que un producto siempre presta un servicio, pero un servicio no necesariamente se materializa en un producto.

Los dos propósitos más comunes por los que se generan productos y servicios de información son la divulgación e intercambio del conocimiento y la puesta a disposición de herramientas que permitan responder preguntas o solucionar problemas.

Las audiencias pueden variar de acuerdo con el grado de complejidad del producto o servicio de información. Se pueden distinguir por lo menos los siguientes niveles de audiencias: 1) científico y académico, 2) tomador de decisiones y 3) público en general (Franco Villegas 2003).

Los productos y servicios constituyen el punto de enlace entre los generadores (*involucrados*) y los usuarios (*interesados y público en general*) de la información (Franco-Villegas 2003). Si se considera que el conocimiento es un nivel agregado de los datos y la información porque permite hacer predicciones, resolver problemas y tomar decisiones, los productos y servicios son una forma de llevar ese conocimiento a una forma concreta.

Es posible definir un conjunto de características deseables en los productos y servicios de información (WCMC 1998b, CADRC 2004). Los productos de información deben:

- Estar dirigidos a audiencias específicas y tener un propósito determinado.
- Basarse en principios científicos y en datos de alta calidad.
- Ser de fácil y rápida comprensión. La interacción usuario-producto se facilita con dos características: un alto nivel de representación de objetos y una interfaz intuitiva (CADRC 2004). La interfaz de usuario debe ser gráfica por naturaleza. En conjunto, el producto debe ser fácilmente operable, de modo que los usuarios puedan aprender a utilizarlo de forma autónoma. Sin embargo, siempre debe estar disponible un sistema de ayuda.
- Estar acompañados por un reconocimiento completo de las fuentes de información y de la propiedad intelectual.
- Ser relevantes y oportunos para necesidades de toma de decisiones.
- Divulgarse a través de canales reconocidos.
- Estar disponibles a costos mínimos de tiempo, dinero y administración.
- Tener afinidad con referentes nacionales e internacionales.

Los problemas complejos usualmente involucran actores que pueden estar geográficamente dispersos y que utilizan recursos de información igualmente dispersos. Por lo tanto, cuando se requiera la operación distribuida, es necesario tener en cuenta en el diseño todas las consideraciones de estructuración lógica e infraestructura, para garantizar la interoperabilidad de los datos y de los productos.

En muchos casos puede ser deseable diseñar esquemas descentralizados de operación. En un esquema descentralizado las solicitudes de servicio y negociaciones involucran sólo unos pocos nodos en el sistema, por lo que se reduce la probabilidad de formación de cuellos de botella y se aumenta la cantidad de actividad paralela (CADRC 2004). Esta capacidad de combinar componentes semiautónomos, que representan un amplio rango de intereses y que incorporan diferentes clases de conocimiento, proporciona una gran versatilidad y potencial para resolver problemas.

Un *producto* usualmente incorpora una gran base de conocimiento implícito. Este conocimiento ayuda pero no reemplaza la capacidad de tomar decisiones. Las capacidades humanas de conceptualización, intuición y creatividad son irremplazables, pero se complementan con fortalezas de un producto informático tales como velocidad, paralelismo, exactitud y capacidad de almacenamiento.

Por lo anterior, los productos deben concebirse como herramientas que permiten llegar a la solución de un conjunto de problemas, pero no como la solución misma. La naturaleza indeterminada de los problemas complejos no permite predecir las circunstancias específicas de una futura situación problema o los términos precisos de la solución. Es más útil tener herramientas que puedan usarse en un amplio rango de situaciones, que herramientas demasiado específicas y circunscritas a espectros de análisis muy pequeños.

a. Tipos de productos según el tiempo de desarrollo requerido

De acuerdo con los tiempos de desarrollo se pueden definir tres categorías de productos de información (WCMC 1998b; Dávila 2003, citado en Franco Villegas 2003):

1. *Productos de desarrollo inmediato (<3 meses)*: el desarrollo del producto se puede iniciar tan pronto como los recursos sean asignados, pues los conjuntos de datos básicos ya existen y son de buena calidad.
2. *Productos de desarrollo intermedio (3 meses a 1 año)*: se requiere compilar conjuntos de datos, para lo cual se necesita de la participación de expertos en la materia.
3. *Productos de desarrollo estratégico (> a 2 años)*: los conjuntos de datos básicos para la conformación del producto no existen, son inaccesibles o no cumplen con estándares que aseguren su calidad, por lo cual es necesario hacer el levantamiento de la información.

b. Etapas del desarrollo de un producto

A continuación se presenta una secuencia de los pasos para desarrollar un producto de información típico. No todos los pasos aplican para todos los tipos de productos; los señalados con (*) son requerimientos comunes para casi todos los productos.

- *1. Definición del problema que el producto pretende abordar: propósito concreto del producto, qué pregunta responde, qué problema contribuye a resolver.
- *2. Conceptualización del problema y de la respuesta. Definición de el(los) nivel(es) de la biodiversidad que aborda (ecosistemas, comunidades, organismos, genes).

En algunas ocasiones, el desarrollo de un producto requiere tomar decisiones sobre bases conceptuales que pueden ser contradictorias. Cuando esto ocurra, la decisión debe documentarse y argumentarse apropiadamente.

- *3. Definición de la forma como se espera que el producto entregue la respuesta. Algunas opciones son: valores individuales, listas provenientes de búsquedas, tablas, gráficas, mapas, resultados de búsquedas distribuidas, entre otros.

Con estos tres pasos se puede obtener un mapa mental claro del producto y una visión preliminar de usos, usuarios, nivel tecnológico e insumos requeridos.

- *4. Revisión de productos existentes que cumplan funciones iguales o similares a las requeridas. En la revisión se deben tener en cuenta aspectos de estructura de la información, ambientes de desarrollo y funcionalidad.
- *5. ¿Se necesita desarrollar un producto de información o ya existe uno que puede usarse para el propósito del problema planteado?

Uno de los factores de decisión es la evaluación del impacto del producto en términos de utilidad, magnitud del problema que aborda, usuarios, novedad y relación costo-beneficio entre costos de desarrollo e implementación *vs.* impacto. Si se opta por desarrollar un producto, éste debe tener afinidad con los referentes nacionales e internacionales reconocidos.

- *6. Caracterización de usuarios. Los aspectos que deben tenerse en cuenta son:
 - número de usuarios potenciales
 - *hardware* y *software* disponibles
 - capacidad humana
 - acceso a Internet
 - sitios y condiciones de uso del producto (campo u oficina, computador, servidor, red LAN, red WAN, etc.)
 - posibilidad futura de aumentar su capacidad actual
 - voluntad de involucrarse en un proceso de implementación de productos
- *7. Definición del alcance técnico del producto. Si se decide que se va a desarrollar un producto, es necesario acotar su alcance en términos de funcionalidad. Para esto se debe considerar todo el panorama de funcionalidades posibles que se podrían implementar con la tecnología existente, es decir, contemplar todo el horizonte ideal de desarrollo; sobre esa base se establece un nivel mínimo de desarrollo, consistente con los recursos, el tiempo y la información disponibles. La consideración de ese panorama de funcionalidad es importante para hacer las previsiones pertinentes en diseño e infraestructura, de manera que el producto desarrollado pueda ser susceptible de mejoramientos futuros.
- *8. Levantamiento y descripción detallada de requerimientos. Es importante establecer si se requieren funciones de análisis espacial. También se deben establecer las necesidades de conjuntos de datos de referencia.
- *9. Revisión de fuentes y condiciones de los datos. Se debe contemplar:
 - formato y medio físico en el que se encuentran
 - estructura
 - tamaño del conjunto de datos

- calidad y confiabilidad de los datos
- condiciones de acceso y uso
- necesidades de acuerdos para tener acceso a los datos
- *10. Definición de necesidades de procesamiento de datos.
- *11. Revisión de necesidades de interoperabilidad con datos o productos existentes.
- 12. Definición de infraestructura de desarrollo: infraestructura física, tecnología (motores de bases de datos, lenguajes de programación, protocolos de intercambio de información) y procedimientos de seguridad.
- 13. Diseño y construcción de la base de datos.
- 14. Pruebas de la base de datos.
- 15. Desarrollo de la aplicación. Se incluyen en este punto: diseño del mapa de navegación, diseño gráfico de la interfaz, desarrollo del núcleo de la aplicación, implementación de niveles de control de acceso y modificación de los datos, implementación de mecanismos de control de errores.
- 16. Elaboración de consultas a la base de datos.
- 17. Pruebas de la aplicación.
- *18. Revisión y pruebas con usuarios potenciales.
- 19. Documentación de la base de datos y de la aplicación, y elaboración de manuales de usuario y de administrador.
- *20. Distribución e instalación del producto.
- 21. Capacitación a usuarios.
- 22. Soporte y mantenimiento.

a. Criterios para la evaluación de productos y servicios de información

- Caracterización de audiencia y objetivos
- Cobertura y alcance del contenido
- Estructura lógica
- Claridad del diseño y la interfaz
- Condiciones de accesibilidad
- Disponibilidad oportuna de datos e información
- Opciones para la extracción de datos e información
- Frecuencia de mantenimiento y actualización de la información
- Relevancia para la toma de decisiones
- Relación costo/beneficio entre desarrollo e impacto del producto
- Experiencia y trayectoria del autor
- Reconocimiento completo de las fuentes de información y de la propiedad intelectual
- Retroalimentación con usuarios
- Atención a usuarios
- Afinidad con referentes nacionales o internacionales.

XI. Literatura citada

Bello J. C., Leyva P., Andrade-C M. G., Calle S. 2003. Hacia la implementación de un sistema de información sobre biodiversidad para Colombia. Versión 1.1 MS.

Berendsohn W. G. 1997. A taxonomic information model for botanical databases: the IOPI Model. *Taxon* 46: 283-309. URL: http://www.bgbm.org/BioDivInf/Docs/IOPI_Model/ Fecha de actualización: 2003. [Fecha de consulta: 20040308]

Berendsohn W. G., Anagnostopoulos A., Hagedorn G., Jakupovic J., Nimis P. L., Valdés P., Güntsch A., Pankhurst R. J. & White R. J. 1999. A comprehensive reference model for biological collections and surveys. *Taxon* 48: 511 - 562.

BCIS - Biodiversity Conservation Information System. 2000a. Framework for information sharing: Custodianship. Busby J. R. (ed.). URL:). <http://www.biodiversity.org>

BCIS - Biodiversity Conservation Information System. 2000b. Framework for information sharing: Principles Busby J. R. (ed.). URL:). <http://www.biodiversity.org>

BCIS - Biodiversity Conservation Information System. 2000c. Framework for information sharing: Data access Busby J. R. (ed.). URL:). <http://www.biodiversity.org>

Bisby F. A. 1994. Plant names in botanical databases. Plant taxonomic database standards No. 3. Versión 1.00. International Working Group on Taxonomic Databases for Plant Sciences (TDWG).

Blum S. D. 1998. A workshop on the compilation, maintenance, and dissemination of taxonomic authority files. Propuesta sobre el taller presentada a la National Science Foundation. Documento electrónico. URL: http://research.calacademy.org/taf/TAF_NSF_Proposal.PDF [Fecha de consulta: 20040308]

Boyle B.L. 20040921. TaxonScrubber Version 1.2. The SALVIAS Project, <http://www.salvias.net>. [F. consulta: 20061110].

CADRC - Collaborative Agent Design Research Center 2004. Decision-Support Systems. URL: http://www.cadrc.calpoly.edu/pdf/decision_brochure.pdf. [F. consulta: 20040210]

Decreto 1603 de 1994 (julio 27). Por el cual se organizan y establecen los Institutos de Investigación de Recursos Biológicos “Alexander Von Humboldt”, el Instituto Amazónico de Investigaciones “SINCHI” y el Instituto de Investigaciones Ambientales del Pacífico “John Von Neumann”. Bogotá, Colombia.

Decreto 2811 de 1974 (diciembre 18). Por el cual se dicta el Código Nacional de Recursos Naturales Renovables y de Protección al Medio Ambiente. Bogotá, Colombia.

- Escalante T., Llorente J., Espinosa D., Soberón J. 2000. Bases de datos y sistemas de información: Aplicaciones en biogeografía. *Rev. Acad. Colomb. Cienc.* 24 (92): 325-341.
- FGDC - Federal Geographic Data Committee. 2000. Content Standard for Digital Geospatial Metadata Workbook Version 2. Washington, D. C.
- Franco Villegas X. 2003. Política institucional para la gestión de datos e información sobre biodiversidad. Versión 1.3. (electrónica). Instituto Alexander von Humboldt. Bogotá D.C., Colombia, 65 pp.
- Franco Villegas X. 2003. Política institucional para la gestión de información sobre biodiversidad. Versión 1.2 (electrónica). Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, D.C, Colombia, 60 p.
- Franco Villegas X. 2004. Esquema general de membresías del Sistema de Información sobre Biodiversidad. Versión 2. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, D.C, Colombia, 6 p.
- Franco Villegas X., Bello J. C. 2004. Estrategia de gestión interinstitucional para la implementación del Sistema de Información sobre Biodiversidad. Documento interno. Tercera versión. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, D.C., Colombia, 11 p.
- French J. C., Jones A. K., Pfaltz J. L. 1990. Scientific database management (final report). Technical report 90-21, Department of Computer Science, University of Virginia.
- Gaston, K. J. (ed.). 1996. Biodiversity: a biology of numbers and difference. Blackwell Science Ltd.
- Greuter W., McNeill J., Barrie F. R., Burdet H. M., Demoulin V., Filgueiras T. S., Nicolson D. H., Silva P. C., Skog J. E., Trehane P., Turland N. J., Hawksworth D. L. (eds.). 2000. International code of botanical nomenclature (Saint Louis Code). Documento electrónico. URL:<http://www.bgbm.fu-berlin.de/iapt/nomenclature/code/SaintLouis/0001ICSLContents.Htm> [Fecha de consulta: 20040305]
- ICZN - International Commission on Zoological Nomenclature. 2000. International code of zoological nomenclature. 4ª edición.
- ITIS - Integrated Taxonomic Information System. 2001. Guidelines for submitting names & associated data to ITIS. Base de datos en línea. URL: http://www.itis.usda.gov/submit_guidelines.pdf. [Fecha de consulta: 20040308]
- ISO – International Organization for Standardization. 2007. ISO 9000 and ISO 14000 - in brief. URL: <http://www.iso.ch/iso/en/iso9000-14000/understand/inbrief.html> [Fecha de consulta: 20070630]
- IUCN Species Survival Commission. 2001. IUCN Red List Categories, version 3.1. IUCN SSC. 24 p.
- Ley 99 de 1993 (diciembre 22). Por la cual se crea el Ministerio del Medio Ambiente, se reordena el sector público encargado de la gestión y conservación del medio ambiente y los recursos naturales renovables, se organiza el Sistema Nacional Ambiental (SINA) y se dictan otras disposiciones. Bogotá, Colombia.
- Mayorga Rodríguez C. 2002. Metodología de la investigación. Editorial Panamericana, Bogotá, 134 p.
- Michener W. K., Brunt J. W., Helly J. J., Kirchner T. B., Stanfford S. G. 1997. Nongeospatial metadata for ecological sciences. *Ecological Applications*, 7(1): 330-342. Ecological Society of America.
- Noss R. F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4 (4): 355-364.
- Real Academia Española de la Lengua. 2001. Diccionario de la lengua española, XXI edición, electrónica. URL: www.rae.es. [Fecha de consulta: 20030301]

Rivera-Gutiérrez H. F. & Suárez-Mayorga A. M. 2005. Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad, versión 2.0. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D.C., Colombia, 66 p.

Rivera-Gutiérrez H. F., Suárez-Mayorga A. M., Varón-Londoño A. 2003a. Estándar para la documentación de metadatos de conjuntos de datos relacionados con biodiversidad. Versión electrónica. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia. 66 p.

Rivera-Gutiérrez H. F., Suárez-Mayorga A. M., Bello-Silva J. C. 2003b. Estándar para la documentación de registros biológicos. Versión 4.1 (electrónica). Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia, 63 p.

Stenmark D. 2002. Information vs. knowledge: The role of intranets in knowledge management. Proceedings of the 35 Hawaii International Conference on System Science 1-10.

Suárez-Mayorga A. M., Rivera-Gutiérrez H. F., Varón-Londoño A., Ramón N. A. 2005a. Estándar para la documentación de registros biológicos, versión 5.0. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D. C., Colombia, 64 p.

Suárez-Mayorga A. M., Bernal R. G., Cárdenas-López D. 2005b. Estándar para intercambiar información al nivel de organismos. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá D. C., Colombia, 23 p.

Suárez-Mayorga A. M. 2003. Documento explicativo del catálogo de métodos y atributos. Versión 1.0 (preliminar). Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia. 8 p.

Trujillo-Motta D., Varón-Londoño A, Bello-Silva J.C. 2005. Estándar para la elaboración de archivos de autoridad taxonómica. Versión 5. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá DC, Colombia, 51 p.

